Titre: Title:	Differentially Private Event Stream Filtering with an Application to Traffic Estimation
Auteur: Author:	Meisam Mohammady
Date:	2015
Type:	Mémoire ou thèse / Dissertation or Thesis
Référence: Citation:	Mohammady, M. (2015). Differentially Private Event Stream Filtering with an Application to Traffic Estimation [Master's thesis, École Polytechnique de Montréal]. PolyPublie. https://publications.polymtl.ca/1739/

UNIVERSITÉ DE MONTRÉAL

DIFFERENTIALLY PRIVATE EVENT STREAM FILTERING WITH AN APPLICATION TO TRAFFIC ESTIMATION

MEISAM MOHAMMADY
DÉPARTEMENT DE GÉNIE ÉLECTRIQUE
ÉCOLE POLYTECHNIQUE DE MONTRÉAL

MÉMOIRE PRÉSENTÉ EN VUE DE L'OBTENTION
DU DIPLÔME DE MAÎTRISE ÈS SCIENCES APPLIQUÉES
(GÉNIE ÉLECTRIQUE)
AVRIL 2015

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Ce mémoire intitulé :

DIFFERENTIALLY PRIVATE EVENT STREAM FILTERING WITH AN APPLICATION TO TRAFFIC ESTIMATION

présenté par : MOHAMMADY Meisam

en vue de l'obtention du diplôme de : Maîtrise ès sciences appliquées

a été dûment accepté par le jury d'examen constitué de :

M. MALHAMÉ Roland, Doct., président

M. LE NY Jérôme, Ph. D., membre et directeur de recherche

M. FAROOQ Bilal, Ph.D., membre

DEDICATION

To my mum...

ACKNOWLEDGEMENT

I would like to thank Jerome Le Ny, my supervisor and my friends Mohsen Ghafouri and Moein Karami for their supports during my Masters program. I also must thank Rabih Salhab for his kind help in translating the abstract to french.

RÉSUMÉ

Beaucoup de systèmes à grande échelle tels que les systèmes de transport intelligents, les réseaux intelligents ou les bâtiments intelligents requièrent que des individus contribuent leurs flux de données privées afin d'amasser, stocker, manipuler et analyser les informations pour le traitement du signal et à des fins de prise de décision. Dans un scénario typique, un essaim de capteurs produit des signaux d'entrée à valeurs discrètes décrivant l'occurrence d'événements relatifs à ces individus. En conséquence, des statistiques utiles doivent être publiées continuellement et en temps réel. Cependant, cela peut engendrer une perte de confidentialité pour les utilisateurs. Cette thèse considère le problème de fournir des garanties de confidentialité différentielle pour ces systèmes multi-sorties multi-entrées fonctionnant en continu. En particulier, nous considérons la question de confidentialité dans le contexte de la théorie des systèmes et nous étudions le problème de génération de signaux qui respectent la confidentialité des utilisateurs qui activent les capteurs. Nous présentons une nouvelle architecture d'estimation des flux de trafic préservant la confidentialité des conducteurs. Nous introduisons aussi une surveillance différentiellement confidentielle d'occupation dans un bâtiment équipé d'un dense réseau de capteurs de détection de mouvement, qui sera utile par exemple pour commander le système HVAC.

ABSTRACT

Many large-scale systems such as intelligent transportation systems, smart grids or smart buildings require individuals to contribute their private data streams in order to amass, store, manipulate and analyze information for signal processing and decision-making purposes. In a typical scenario, swarms of sensors produce discrete-valued input signals that describe the occurrence of events involving these users and several statistics of interest need to be continuously published in real-time. This can however engender a privacy loss for the users in exchange of the utility provided by the application. This thesis considers the problem of providing differential privacy guarantees for such multi-input multi-output systems operating continuously. In particular, we consider the privacy issues in a system theoretic context, and address the problem of releasing filtered signals that respect the privacy of users who activate the sensors. As a result of this thesis we present a new architecture for privacy preserving estimation of traffic flows. We also introduce differentially private monitoring and forecasting occupancy in a building equipped with a dense network of motion detection sensors, which is useful for example to control its HVAC system.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ACRONYMS AND ABBREVIATIONS

SISO	Single-Input Single-Output
SIMO	Single-Input Multiple-Output
MISO	Multiple-Input Single-Output
MIMO	Multiple-Input Multiple-Output
CTM	Cell Transmission Model
EKF	Extended Kalman Filter
i.i.d.	Identically and Independently Ditributed

CHAPITRE 1 INTRODUCTION

Recently, applications which require individuals to contribute their private data in order to amass, store, manipulate and analyze information for decision-making purposes have become increasingly popular. But, whether this aggregated data is used in the service of a specific business, political party or even a new scientific finding, researchers also raise questions for the individuals whose information comprises these "big data" sets. Is it really possible to guarantee that the individuals' information will remain private ?
Before start talking about different anonymization techniques, we first recap two famous information breaches that have occurred in two large companies (AOL and Netflix) and have been widely covered both in the academic literature and news media.

1.1 AOL Search Log Leakage

On August 2nd, 2006, the AOL search log released a file consisting more than 20 million search queries for over 600,000 users over a period of 3-months. This file was released for research purposes and for security issues AOL deleted the search data on their site on next day. However the file circulated widely on internet and was even restored by mirror websites[4]. The information amassed in the report was anonymized and did not re-identify the users, however, some personal information presented in many of the keywords could be linked with external information in order to uniquely identify individuals. For example, the New York Times published on August 9, 2006, successfully discerned the user No. 4417749, as follows [5]. "No. 4417749 conducted hundreds of searches over a three-month period on topics ranging from "numb fingers" to "60 single men" to "dog" that urinates on everything. It did not take much investigating to follow that data trail to Thelma Arnold, a 62-year-old widow who lives in Lilburn, Ga."

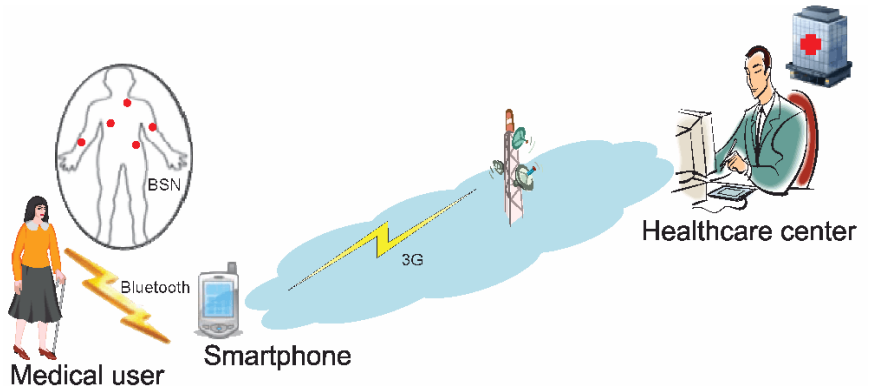
1.2 Netflix Prize

Netflix, Inc. is an online service provider of movie streaming available to viewers in many countries. In October, 2006, Netflix released a dataset that included over 100 million ratings contributed by over 450 thousand users to 18000 movies. They announced a contest among the data mining, machine learning and computer science communities with a million dollar prize for the best prediction algorithm that could beat their existing movie recommendation algorithm[6]. The dataset provided by Netflix however was modified deliberately to protect the privacy of the users. For

example, they removed all sensitive information such as name, user-name, age, and location[7]. They even intentionally perturbed "some of the ratings" for random customers in one of the following ways : deleting some of their ratings, and modifying the rating dates [6]. The released data consisted of attributes such as randomly assigned numeric user id, movie, date and the value of the rating on a particular scale. However, one year later A. Narayanan and V. Shmatikov [8] showed how to re-identify several users in the released Netflix file simply by crosscorrelating the anonymized ratings with non-private movie ratings on the Internet Movie Database (IMDb) website.
Basically, linkage attacks aim at linking the output of a query to a record or a value in a given table to establish the presence or absence of a target individual. These type of attacks are divided into two major groups of deterministic attacks and stochastic attacks [1].

1.3 Deterministic linkage attacks

These attacks mainly try to link an individual to an unprotected or deterministically protected published dataset based on arbitrary side information, including user level information,(name, sex,etc) or an exclusive value that belongs to the victim like her carrier. Deterministic methods of sanitizing a dataset mainly rely on anonymizing, e.g, Sweeney [9] proposed k-anonymity which requires that for each set of attributes in the table, there must be at least $k-1$ other records with the same attribute. For example, a health care service center who frequently updates its patients' symptoms, can be attacked by adversaries as shown in Fig. 1.1 . The information sorted in the table has already been anonymized, however the level of anonymity is not enough to prevent information leakage of a particular patient, named Audrey. By increasing the level of anonymity to 3-anonymity as shown in Fig. 1.2 , the published table becomes protected against this particular type of attacks.
However, even Table . 1.2 could be vulnerable to some attacks, for example, assume the attacker knows Audrey and he also knows that she does not have the visible flu symptoms. According to the data Table . 1.2, he could infer Audrey has HIV.



Job	Sex	Age	Disease
Pianist	male	35	Hepatitis
Actor	male	38	Hepatitis
Teacher	male	38	HIV
Actress	female	35	Flu
Writer	female	35	HIV
Dancer	female	35	HIV
Dancer	female	36	HIV

Anonymized Patients table

Audrey

Name	Job	Sex	Age
Audrey	Waitress	Female	31
...	...	...	...

Attacker's side attributes

Figure 1.1 Anonymized Table of attributes for each patient recorded by healthcare center could be linked with side attributes available to an attacker. For example, although Audrey's record has already been anonymized, an attacker who knows that Audrey is under treatment could identify her disease according to the published anonymized patients table.

Job	Sex	Age	Disease
Professional	male	[35-40)	Hepatitis
Professional	male	[35-40)	Hepatitis
Professional	male	[35-40)	HIV
Actress	female	[35-40)	Flu
Actress	female	[35-40)	HIV
Actress	female	[35-40)	HIV
Actress	female	[35-40)	HIV

Figure 1.2 3-anonymous table [1]

The main drawback of this notion of privacy is the privacy preserving method does not model the adversaries' background knowledge. To resolve this problem, stochastic privacy preserving models

popped up.

1.4 Stochastic linkage attacks

These attacks mainly try to conclude a particular individual's attribute in a probabilistic way. For example, assume that an adversary knows the victim has HIV with probability 0.2. If he can derive from the published table that the victim has HIV with probability 0.7, then he has done a probabilistic attack successfully, since the difference between the prior (0.2) and posterior (0.7) probabilities is meaningfully big [1]. On the contrary, stochastic privacy models that deal with probabilistic attacks often try to restrict the difference between an adversary's prior and posterior probabilistic knowledge of an attribute of a victim. All of these techniques rely on randomization of the database and each model has designed to handle a particular type of attack.

1.5 Objectives of the thesis

In the last few years, a strong notion of privacy, namely, *differential privacy* has emerged, which provides some privacy guarantees against adversaries with arbitrary side information. Differential Privacy [10] aims at limiting the risk enhancement to one's privacy when she contributes her data to a statistical database [7]. This model ensures that adding or removing a single record does not significantly affect the outcome of the sanitized algorithm. In this thesis, we focus on the notion of (ϵ, δ)-differential privacy that formalizes the trade-off between privacy and utility in sanitizing an applications. The smaller values of parameters ϵ and δ indicate the stronger privacy guarantees provided to the users' records. However strong privacy guarantees could have negative impacts on the performance of the applications analyzing these records. In other words, the stronger privacy guarantee for an application outcome requires the bigger noise injection to the database. Hence, differential privacy could require large perturbations to an analysis outcome in order to hide the presence of individuals [11] . This is especially true for applications where users frequently contribute data over time, and it is hence crucial to design advanced mechanisms restricting the impact on system performance of differential privacy constraints [11] . These applications could vary from intelligent transportation systems to smart grids and smart buildings. This thesis focuses on designing a privacy preserving architecture for such a real-time applications from a signal processing perspective. Previous work on designing differentially private time-series mechanisms includes [12] and [13], where they propose structures that could not be implemented in real-time. Inspired by [11], here, we consider the problem of designing differentially private real-time architectures for time-series applications. We extend the mechanism of [11] which was proposed for Single-Input Single-Output applications, to the Multiple-Input Multiple-Output (MIMO) type of time-series queries. Our result

considerably broadens the applicability of the ideas to real-time applications compared to the analysis in [11]. The rest of the thesis is organized as follows. We first present in Chapter 2, necessary theoretical background on differential privacy. We then consider the problem of providing differential privacy guarantees for MIMO systems in Chapter 3 [14] The theoretical analysis results are then implemented in a real world application which aims at privately estimating and forecasting occupancy in a building equipped with a network of motion detectors. The goal of a differential privacy constraint in the building monitoring application is that an individual user cannot be tracked too precisely from the published data. In Chapter 4 , we present a differentially private real time traffic estimator. Our results improves over the state of the art and has immediate applications in Highway Traffic Management Systems (HTMS). Finally, we conclude with a brief review in Chapter. 5 .

CHAPITRE 2 DIFFERENTIAL PRIVACY

In this chapter we review the notion of differential privacy which provides strong quantitative privacy guarantees for the users data. As mentioned in the introduction 1.4, differential privacy is a randomization-based notion of privacy. This notion could securely compose even in the presence of arbitrary side information [**?**]. This model ensures that adding or removing a single user's data record does not significantly affect the outcome of any data synthesis [15].

Definition of Differential Privacy

Formally, we start by defining a symmetric binary relation, denoted Adj, on the space of datasets D of interest, which is used to define what it means for two datasets to differ by the data of a single individual. For any d,d' subsets of D , we have Adj(d, d') if and only if we can obtain the signal d' from d simply by adding or subtracting the data of one user. Then mechanisms that are differentially private necessarily randomize their outputs, in such a way that they satisfy the following property.

Definition 1. *Let* D *be a space equipped with a symmetric binary relation denoted Adj, and let* $(\mathsf{R}, \mathcal{M})$ *be a measurable space. Let* $\epsilon, \delta \geq 0$*. A mechanism* $M : \mathsf{D} \times \Omega \to \mathsf{R}$ *is* (ϵ, δ)*-differentially private for Adj if for all* $d, d' \in \mathsf{D}$ *such that Adj*(d, d')*, we have*

$$\mathbb{P}(M(d) \in S) \leq e^{\epsilon}\mathbb{P}(M(d') \in S) + \delta, \quad \forall S \in \mathcal{M}. \tag{2.1}$$

If $\delta = 0$*, the mechanism is said to be* ϵ*-differentially private.*

This definition quantifies the allowed deviation for the output distribution of a differentially private mechanism, when a single individual is added or removed from a dataset. If the inequality fails, then a leakage ((ϵ, δ) breach) takes place. This simply means that the difference between the prior distribution and posterior one is tangible.

The choice of the parameters ϵ, δ is set by the privacy policy. Typically ϵ is taken to be a small constant, e.g., $\epsilon \approx 0.5$ or perhaps even $\ln p$ for some small $p \in \mathbb{N}$. The parameter δ should be kept small as it controls the probability of certain significant losses of privacy, e.g., when a zero probability event for input d' becomes an event with positive probability for input d in (2.1). One fundamental property of the notion of differential privacy which is widely used in Chapter 3 , is that no additional privacy loss can occur by simply manipulating an output that is differentially private. To state it, recall that a probability kernel between two measurable spaces $(\mathsf{R}_1, \mathcal{M}_1)$ and $(\mathsf{R}_2, \mathcal{M}_2)$ is a function $k : \mathsf{R}_1 \times \mathcal{M}_2 \to [0, 1]$ such that $k(\cdot, S)$ is measurable for each $S \in \mathcal{M}_2$ and $k(r, \cdot)$ is a probability measure for each $r \in \mathsf{R}_1$.

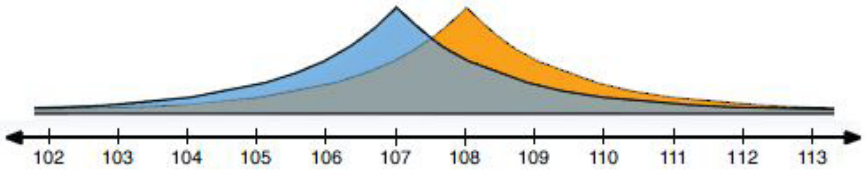


Figure 2.1 Probability distributions on two Adjacent data bases

Theorem 1 (Resilience to post-processing [16]). *Let* $M_1 : \mathsf{D} \times \Omega \to (\mathsf{R}_1, \mathcal{M}_1)$ *be an* (ϵ, δ)*-differentially private mechanism. Let* $M_2 : \mathsf{D} \times \Omega \to (\mathsf{R}_2, \mathcal{M}_2)$ *be another mechanism, such that there exists a probability kernel* $k : \mathsf{R}_1 \times \mathcal{M}_2 \to [0, 1]$ *verifying*

$$\mathbb{P}(M_2(d) \in S | M_1(d)) = k(M_1(d), S), \text{ almost surely}, \tag{2.2}$$

for all $S \in \mathcal{M}_2$ *and* $d \in \mathsf{D}$*. Then* M_2 *is* (ϵ, δ)*-differentially private.*

Note that in (2.2), the kernel k is not allowed to depend on the dataset d. In other words, this condition says that once $M_1(d)$ is known, the distribution of $M_2(d)$ does not further depend on d. Hence a mechanism M_2 accessing a dataset only indirectly via the output of a differentially private mechanism M_1 cannot weaken the privacy guarantee. Another feature of differential privacy that is used in the thesis is the characterization of differential privacy under adaptive composition. The following theorem shows that the privacy degrades under composition to the sum of the differential privacy parameters of each access.

Theorem 2 ([17]). *Consider* $M_1, ...M_r$*, r mechanisms on a space* D*, where* M_i *is* (ϵ_i, δ_i)*-differentially private. Then the mechanism* $M = (M_1, ..., M_r)$*, which, for* $d \in D$*, outputs* $(M_1(d), ...M_r(d))$*, is at least* $(\sum_{i=1}^{r} \epsilon_i, \sum_{i=1}^{r} \delta_i)$*-differentially private.*

However, it is noted that differential privacy might return answers to queries that may not be useful in practice. A mechanism that throws away all the information in a dataset is obviously private, but not useful, and in general one has to trade off privacy for utility when answering specific queries. We are only concerned in this thesis with queries that return numerical answers, i.e., here a query is a map $q : \mathsf{D} \to \mathbb{R}$ is equipped with an euclidean norm denoted $\|\cdot\|_{\mathbb{R}}$, and the σ-algebra $\mathcal{M}$ on $\mathbb{R}$ is taken to be the standard Borel σ-algebra. The following quantity plays an important role in the design of differentially private mechanisms [18].

Definition 2. *Let* D *be a space equipped with an adjacency relation Adj. The sensitivity of a query* $q : \mathsf{D} \to \mathsf{R}$ *is defined as* $\Delta_{\mathsf{R}} q := \max_{d,d' : Adj(d,d')} \|q(d) - q(d')\|_{\mathsf{R}}$*. In particular, for* $\mathsf{R} = \mathbb{R}^k$ *equipped with the p-norm* $\|x\|_p = \left(\sum_{i=1}^{k} |x_i|^p\right)^{1/p}$*, for* $p \in [1, \infty]$*, we denote the* ℓ_p *sensitivity by* $\Delta_p q$*.*

The magnitude of the distortion required for differentially private publishing the outcomes of a query is then demonstrated in the following theorem. We recall below a basic mechanism that is proposed in [19], perturbs the numerical outcomes of a query by adding i.i.d. zero-mean Gaussian noise. Recall the definition of the Q-function $Q(x) := \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-\frac{u^2}{2}} du$, we have [16, 19]

Theorem 3. *Let $q : \mathsf{D} \to \mathbb{R}^k$ be a query. Then the Gaussian mechanism $M_q : \mathsf{D} \times \Omega \to \mathbb{R}^k$ defined by $M_q(d) = q(d) + w$, with $w \sim \mathcal{N}\left(0, \sigma^2 I_k\right)$, where $\sigma \geq \frac{\Delta_2 q}{2\epsilon}(K + \sqrt{K^2 + 2\epsilon})$ and $K = Q^{-1}(\delta)$, is (ϵ, δ)-differentially private.*

For the rest of the thesis, we define $\kappa_{\delta,\epsilon} = \frac{1}{2\epsilon}(K + \sqrt{K^2 + 2\epsilon})$, so that the standard deviation σ in Theorem 3 can be written $\sigma(\delta, \epsilon) = \kappa_{\delta,\epsilon}\Delta_2 q$. The mechanism M described in Theorem 3, producing a differentially-private version of a query q, is called an output-perturbation mechanism.
To illustrate the definitions and theorems presented in this chapter, let's consider the following example. Assume we have a medical database D_1 where each record is a pair (Name,X), where X $\in \{0, 1\}$, denotes whether a person has HIV or not. For instance

Name	X
Joe	0
Philip	1
Tom	1
Judy	1
Chris	0

Now assume that an attacker wants to find out if Tom has HIV or not. Suppose that he also knows Tom receives his blood test at time t_0. If the health center continuously updates the number of people who has HIV by a query $q(t)$, the adversary simply executes $q(t_0^+) - q(t_0^-)$ and finds Tom's HIV status. This example shows that user's private data could be revealed even without explicitly querying for one particular person's record. Now let's first construct the adjacent dataset D_2 by replacing (Tom,1) with (Tom,0). Based on Definition 2, the sensitivity of q is 1 and in order for (ϵ, δ)-differentially private publishing the outcome of $q(t)$, based on Theorem 3, we must inject i.i.d. Gaussian noise $w \sim \mathcal{N}\left(0, \sigma^2\right)$ to each outcome of $q(t)$, where $\sigma \geq \kappa_{\delta,\epsilon}$. In addition Resilience to post-processing indicates that any analysis on the published $q(t)+w$ could not weaken the privacy guarantee.

CHAPITRE 3 DIFFERENTIALLY PRIVATE EVENT STREAM FILTERING

3.1 Introduction

Recently many applications required to be equipped with swarms of sensors and detectors [20], with the goal of immensely improving the efficiency of standard activities and industries, from healthcare to traffic management systems to power grids. At the same time, it is becoming clear that while releasing and sharing large datasets improves the performance of large-scale systems and helps individuals and organizations to have better data analyses and predictions, these systems could pose a risk to our privacy, and more work is needed to provide rigorous ways to tradeoff privacy with utility [21] . This chapter studies such tradeoffs under the notion of differential privacy in the context of real-time systems. The real-time requirement is motivated by the fact that we would like to integrate our filter outputs into real-time information and decision-making systems, e.g., motion detectors in buildings [14] presented in this chapter or loop inductors in traffic information system which is presented in Chapter 4 . Choosing the notion of differential privacy is motivated by the fact that this notion is particularly convenient to pose privacy guarantee for real-time applications, because its requirements could be respected simply by adding noise to the dataset. The rest of the chapter is organized as follows. Section. 3.2 provides some background on the subject of privacy preserving approximation set-ups of real time systems. Section 3.3 introduces our problem formulation. Section 3.4 presents certain preliminary calculations necessary to develop our differentially-private dynamic mechanisms, which is the object of Section 3.5. Finally, in Section 3.6 we briefly describe an application to privately estimating and forecasting occupancy in a building equipped with a network of motion detectors.

3.2 Background

Previous work on designing differentially private mechanisms for the publication of time-series include [13, 22], but these mechanisms are not causal and hence not suited for real-time applications. The papers [23, 24], [25] provide real-time mechanisms to approximate a few specific filters, transforming event streams input into public output streams. For example, [23, 24] consider a private accumulator providing the total number of events that occurred in the past. This chapter is inspired by this scenario, and builds on J. Le Ny et al work on this problem [26, Section IV] [16, Section VI]. Here we extend the analysis presented in [11, 16, 26] , to multi-input multi-output systems, which considerably broadens the applicability of these ideas to common situations where multiple sensors monitor an environment and we wish to concurrently publish several statistics of interest.

A typical application example is that of analyzing spatio-temporal records provided by networks of simple counting sensors [3] .

3.3 Problem Statement

3.3.1 MIMO Event Stream Filtering

We consider a system equipped with m sensors detecting events, with each sensor i producing a discrete-time sequence $\{u_{i,t}\}_{t\geq 0} \in \mathbb{R}$, for $i \in [1, m]$. In a building monitoring scenario for example, described in Section 3.6 , the sensors could be some motion detectors dispensed at various locations which transmit continuously the number of detected events for period t. We denote u the resulting vector valued signal, i.e., $u_t \in \mathbb{R}^m$. The goal is to release a filtered stream Fu for a particular application, where F is a linear time-invariant system, with m inputs and p outputs. The filter actually takes the input signal from the sensors and publishes output signals y of interest, with $y_t \in \mathbb{R}^p$. For example, we might be interested in continuously monitoring the number of people in various parts of the building. In this case, we could design a real-time estimator updating the number of people in each section. Another application includes short- and medium-term occupancy forecasts, in order to optimize the operations of the Heating, Ventilation, and Air Conditioning (HVAC) system.

The problem considered in this chapter consists in replacing the filter F by a system processing the input u and producing a signal $\tilde{y}$ as close as possible to the desired output $y = Fu$ (here, in the mean-squared sense), while providing some privacy guarantees to the users from which the input signals u originate. The privacy requirement is then explained and quantified in the next subsection.

3.3.2 Differential Privacy For MIMO Event Streams

In this chapter our goal is to pose differential privacy guarantees to the scenarios where individuals continuously share their records with a third party in order for receiving some utilities. In our building monitoring application for example, one goal of a privacy constraint could be to provide guarantees that an individual user cannot be tracked too precisely from the published data. Indeed, Wilson and Atkeson [27] for example demonstrate how to track individual users in a building using a network of simple binary sensors such as motion detectors.

Adjacency Relation

We first start by considering the adjacency relation presented in Chapter. 2. Here, $\mathsf{D} := \{u : \mathbb{N} \mapsto \mathbb{R}^m\}$, and we have $\mathrm{Adj}(u, u')$ if and only if we can obtain the signal u' from u simply by adding or

subtracting the events corresponding to one user. That is

$$\begin{aligned} &\mathrm{Adj}^k(u, u') \text{ iff} \\ &\forall i \in [m], \exists t_i \in \mathbb{N}, \alpha_i, \text{ s.t. } u'_i - u_i = \alpha_i \delta_{t_i}, |\alpha_i| \leq k_i. \end{aligned} \tag{3.1}$$

In other words, it is assumed that a single individual can affect each input signal component at a *single time* (here δ_{t_i} denotes the discrete impulse signal with impulse at t_i), and by at most $k_i \in \mathbb{R}_+$. The following notation will be used in the following. We denote by $k \in \mathbb{R}^m$ the vector with components k_i. Also, let $e_i \in \mathbb{R}^m$ be the i^{th} basis vector, i.e., $e_{ij} = \delta_{ij}$, $j = 1, \ldots, m$. Then for two adjacent signals u, u', we have

$$u' - u = \sum_{i=1}^{m} \alpha_i \delta_{t_i} e_i. \tag{3.2}$$

The adjacency definition (3.1) indicates that for an individual to be differentially private protected, he must respect two main constraints. First he can only activate each sensor once in order to report his event. This requirement makes sense in applications like traffic monitoring (Chapter 4) with fixed loop detectors activated only once by each car traveling along a road, or for certain location-based services where a customer would check-in say at most once per day at each visited store. For a building monitoring scenario however, one person could trigger the same motion detector several times over a short period of interest. To solve this problem one simple but efficient idea is to split the data stream of problematic sensors into several successive intervals, each considered as the signal from a new virtual sensor, so that an individual's data is presented only once in each interval. A MIMO mechanism can then treat such dataset and offers privacy guarantees. This results in addressing one of the main issues for the applicability of the model proposed in [23, 24]. However, increasing the number of inputs degrades the privacy guarantees or the output quality that we can provide. Hence in general no privacy guarantee will be offered to users who activate the same sensor too frequently. Second, the magnitude of his contribution to the dataset must be bounded by k_i, but this is not really problematic in applications such as motion detection, where we can typically take $k_i = 1$.

Sensitivity

In order to reduce the impact of differential privacy on the performance of one application, we must evaluate as precisely as possible the amount of noise necessary to make a mechanism differentially private. For this purpose, the following quantity based on the sensitivity Definition 2 plays an important role.

Definition 3. *The ℓ_2-sensitivity of a system G with m inputs and p outputs with respect to the*

adjacency relation Adj is defined by

$$\Delta_2^{m,p} G = \sup_{Adj(u,u')} \|Gu - Gu'\|_2 = \sup_{Adj(u,u')} \|G(u-u')\|_2,$$

where by definition $\|Gv\|_2^2 = \sum_{t=-\infty}^{\infty} |[Gv]_t|^2$ *and* $|x| = \left(\sum_{k=1}^{p} |x_k|^2\right)^{1/2}$ *is also used throughout the thesis to denote the Euclidean norm for x in* $\mathbb{R}^p$ *or* $\mathbb{C}^p$.

Then recalling Theorem 3, we conclude the following theorem.

Theorem 4. *Let G be a system with m inputs and p outputs, and with* ℓ_2*-sensitivity* $\Delta_2^{m,p} G$ *with respect to an adjacency relation Adj. Then the mechanism* $M(u) = Gu + w$*, where w is a p-dimensional Gaussian white noise with covariance matrix* $\kappa_{\delta,\epsilon}^2 (\Delta_2^{m,p} G)^2 I_p$*, is* (ϵ, δ)*-differentially private with respect to Adj.*

We see that the required additive noise to release a differentially private version of signal Gu is proportional to the ℓ_2-sensitivity of the filter and to $\kappa_{\delta,\epsilon}$, which can be shown to behave roughly as $O(\ln(1/\delta))^{1/2}/\epsilon$. Note that we must inject noise to each output proportional to the sensitivity of the whole filter G, even if G was diagonal say, otherwise trivial linkage attacks that simply average a sufficient number of outputs could potentially detect the presence of an individual with high probability. Finally, the differentially private mechanism for our original problem could be obtained by simply adding a sufficient amount of noise proportional to the sensitivity of our desired filter F, to the output of the filter. In next section, we discuss computing the sensitivity of filter F in details.

3.4 Sensitivity Calculations

We first recall the $\mathcal{H}_2$ norm of an LTI system with m inputs which plays an important role for the following sensitivity calculations.

$$\|G\|_2^2 = \sum_{i=1}^{m} \|G\delta_0 e_i\|_2^2 = \frac{1}{2\pi} \int_0^{2\pi} \mathrm{Tr}(G^*(e^{j\omega}) G(e^{j\omega})) d\omega.$$

Note from the frequency domain definition that writing $G(z) = [G_{ij}(z)]_{i,j}$ for the $p \times m$ transfer matrix, we have

$$\|G\|_2^2 = \sum_{i,j} \|G_{ij}\|_2^2.$$

3.4.1 Sensitivity for the SIMO and Diagonal Cases

Generalizing the single-input single-output scenario considered in [26] to the case of a system with $m = 1$ but possibly multiple outputs (SIMO), we have immediately the following theorem.

Theorem 5 (SIMO LTI system)**.** *Let G be a stable LTI system with one input and p outputs. For the adjacency relation (3.1),*

$$\Delta_2^{1,p} G = k_1 \|G\|_2,$$

where $\|G\|_2$ is the $\mathcal{H}_2$ norm of G.

Proof. We have immediately

$$\begin{aligned} \|G(u - u')\|_2^2 &= |\alpha_1|^2 \|G\delta_{t_1}\|_2^2 \\ &\leq k_1^2 \|G\|_2^2, \end{aligned}$$

and the bound is attained if $|\alpha_1| = k_1$. □

For a MIMO system, the case where G is diagonal, i.e., its transfer matrix is
$G(z) = \mathrm{diag}(G_{11}(z), \ldots, G_{mm}(z))$,
also leads to a simple sensitivity computation. Note that in this case, we have $\|G\|_2^2 = \sum_{i=1}^m \|G_{ii}\|^2$.

Theorem 6 (Diagonal LTI system)**.** *Let G be a stable* diagonal *LTI system with m inputs and outputs. For the adjacency relation (3.1),*

$$\Delta_2^{m,m} G = \|GK\|_2 = \left(\sum_{i=1}^m \|k_i G_{ii}\|_2^2 \right)^{1/2},$$

where $K = diag(k_1, \ldots, k_m)$.

Proof. If G is diagonal, then for u and u' adjacent, we have from (3.2)

$$\begin{aligned} \|G(u - u')\|_2^2 &= \left\| \sum_{i=1}^m \alpha_i G\delta_{t_i} e_i \right\|_2 \\ &= \|\mathrm{col}(\alpha_1 g_{11} * \delta_{t_1}, \ldots, \alpha_m g_{mm} * \delta_{t_m})\|_2^2, \end{aligned}$$

where $\mathrm{col}(x_1, \ldots x_m)$ denotes a signal with values in $\mathbb{R}^m$ if each x_i is a scalar signal. Here g_{ii} denotes

the impulse response of G_{ii}. Hence

$$\begin{aligned}\|G(u-u')\|_2^2 &= \sum_{i=1}^{m} \|\alpha_i g_{ii} * \delta_{t_i}\|_2^2 \\ &= \sum_{i=1}^{m} |\alpha_i|^2 \|G_{ii}\|_2^2,\end{aligned}$$

and $|\alpha_i| \leq k_i$, for all i. Again the bound is attained if $|\alpha_i| = k_i$ for all i. □

The sensitivity calculations for MISO or general MIMO systems are no longer straightforward, because the impulses on the various input channels, obtained from the difference of two adjacent signals u, u', could possibly influence any given output. Still, the following result provides a simple bound on the sensitivity which will be used later for mechanism optimization.

3.4.2 Upper and Lower Bound for the general MIMO Case

Theorem 7. *Let G be an LTI system with $m \times p$ transfer matrix $G(z) = [G_1(z), \ldots, G_m(z)]$ (i.e., with columns G_i), such that $\|G\|_2 < \infty$. For the adjacency relation (3.1),*

$$\|GK\|_2 \leq \Delta_2^{m,p} G \leq |k| \|G\|_2,$$

where $K = diag(k_1, \ldots, k_m)$ and $|k| = \left(\sum_{i=1}^{m} k_i^2\right)^{1/2}$.

Proof. We have

$$G(u-u') = \sum_{i=1}^{m} \alpha_i \, G\delta_{t_i} e_i,$$

and moreover $\|G\|_2^2 = \sum_{i=1}^{m} \|G\delta_{t_i} e_i\|_2^2$ by definition. For the upper bound, we can write

$$\begin{aligned}\|G(u-u')\|_2 &= \left\| \sum_{i=1}^{m} \alpha_i \, G\delta_{t_i} e_i \right\|_2 \\ &\leq \sum_{i=1}^{m} |\alpha_i| \|G\delta_{t_i} e_i\|_2 \\ &\leq |k| \left(\sum_{i=1}^{m} \|G\delta_{t_i} e_i\|_2^2 \right)^{1/2},\end{aligned}$$

where the last inequality results from the Cauchy-Schwarz inequality. For the lower bound, let us first take $u' \equiv 0$. Then consider an adjacent signal with a single discrete impulse on input channel i of height k_i at time t_i, $i = 1, \ldots, m$, with $t_1 < t_2 < \ldots < t_m$. Let $\eta > 0$. Since $\|G\|_2 < \infty$, $\|G_i u_i\|_2 < \infty$,

and hence $|(G_i u_i)_t| \to 0$ as $t \to \infty$. Hence by taking $t_{i+1} - t_i$ large enough for each $1 \leq i \leq m-1$, i.e., waiting for the effect of impulse i on the output to be sufficiently small, we can obtain a signal u such that

$$\|Gu\|_2^2 = \left\| \sum_{i=1}^{m} G_i u_i \right\|_2^2 \geq \sum_{i=1}^{m} k_i^2 \|G\delta_{t_i} e_i\|_2^2 - \eta.$$

Since this is true for any $\eta > 0$ and $\|G\delta_{t_i} e_i\|_2^2 = \|G_i\|_2^2$, we get $(\Delta_2^{m,p} G)^2 \geq \|GK\|_2^2 = \sum_{i=1}^{m} k_i^2 \|G_i\|^2$. □

For this case, we assume $k_1 = \ldots = k_m$, the upper bound on the sensitivity is $k_1 \|G\|_2 \sqrt{m}$. The squared sensitivity, which is related to the variance of the required amount of privacy-preserving noise in an output perturbation scheme, scales then linearly with the number of inputs. We can contrast this bound to the situation where G is diagonal, in which case the sensitivity is exactly $k_1 \|G\|_2$ from Theorem 6. Moreover, the following example shows that the upper bound of Theorem 7 cannot be improved for the general MISO or MIMO case.

Example 1. *Consider the MISO system $G(z) = [G_{11}(z), \ldots, G_{1m}(z)]$, with $g_{1i} = \delta_{\tau_i}$ the impulse response of G_{1i}, for some times $\tau_1, \ldots, \tau_m$. Then $\|G\|_2^2 = m$. Now let $u' \equiv 0$ and $u = \sum_{i=1}^{m} \delta_{t_i} e_i$, so that u and u' are adjacent, with $k_1 = \ldots = k_m = 1$, and moreover let us choose the times t_i such that $\tau_i + t_i$ is a constant, i.e., take $t_i = \kappa - \tau_i$ for some $\kappa \geq \max_i\{\tau_i\}$. Then $Gu = \sum_{i=1}^{m} g_{1i} * u_i = m\delta_\kappa$, and so $\|Gu\|_2^2 = m^2$. This shows that the bound of Theorem 7 is tight in this case. Note that this happens because all the events of the signal u influence the output at the same time. Indeed, if the times $\tau_i + t_i$ are all distinct, then we get $\|Gu\|_2^2 = m$.*

3.4.3 Exact solution for the MIMO Case

For completeness, we give in this subsection an exact expression for the sensitivity of a MIMO filter. Let G be a stable LTI system with m inputs and p outputs, and state space representation

$$\begin{aligned} x_{t+1} &= Ax_t + Bu_t \\ y_t &= Cx_t + Du_t. \end{aligned} \tag{3.3}$$

Recall the definition of the observability Gramian P_0, which is the unique positive semi-definite solution of the equation

$$A^T P_0 A - P_0 + C^T C = 0.$$

Let B_i, D_i be the i^{th} column of the matrix B and D respectively, for $i = 1, \dots, m$. Finally, define for $i, j \in \{1, \dots, m\}$, $i \neq j$, and τ in $\mathbb{Z}$

$$S_{ij}^{\tau} = \tag{3.4}$$
$$\begin{cases} B_i^T (A^{\tau-1})^T C^T D_j + B_i^T (A^{\tau})^T P_0 B_j, & \text{if } \tau > 0 \\ D_i^T D_j + B_i^T P_0 B_j, & \text{if } \tau = 0 \\ D_i^T C A^{|\tau|-1} B_j + B_i^T P_0 A^{|\tau|} B_j, & \text{if } \tau < 0. \end{cases}$$

Theorem 8. *Let G be a stable LTI system with m inputs and p outputs, and state space representation (3.3). Then, for the adjacency relation (3.1),*

$$(\Delta_2^{m,p} G)^2 = \|GK\|_2^2 + \sum_{\substack{i,j=1 \\ i \neq j}}^{m} k_i k_j \left(\sup_{t_i, t_j \in \mathbb{N}} \left| S_{ij}^{t_i - t_j} \right| \right). \tag{3.5}$$

Proof. In view of (3.2), we have

$$\Delta_2^{m,p} G = \sup_{|\alpha_i| \leq k_i, t_i \geq 0} \left\| \sum_{i=1}^{m} \alpha_i G \delta_{t_i} e_i \right\|_2 .$$

For $y_i = G\delta_{t_i} e_i$ and $y = \sum_{i=1}^{m} \alpha_i y_i$, we have

$$\begin{aligned} \|y\|_2^2 &= \sum_{t=0}^{\infty} \left| \sum_{i=1}^{m} \alpha_i y_{i,t} \right|^2 \\ &= \sum_{t=0}^{\infty} \sum_{i=1}^{m} \alpha_i^2 |y_{i,t}|^2 + \sum_{t=0}^{\infty} \sum_{\substack{i,j=1 \\ i \neq j}}^{m} \alpha_i \alpha_j y_{i,t}^T y_{j,t} \\ &\leq \|GK\|_2^2 + \sum_{\substack{i,j=1 \\ i \neq j}}^{m} k_i k_j \left| \sum_{t=0}^{\infty} y_{i,t}^T y_{j,t} \right|, \end{aligned}$$

where $K = \operatorname{diag}(k_1, \dots, k_m)$ and the bound can be attained by taking $\alpha_i \in \{-k_i, k_i\}$, depending on the sign of $S_{ij} := \sum_{t=0}^{\infty} y_{i,t}^T y_{j,t}$.

Next, we derive the more explicit expression of S_{ij} given in the theorem. First,

$$y_{i,t} = \begin{cases} 0, & t < t_i, \\ D_i, & t = t_i \\ C A^{t - t_i - 1} B_i, & t > t_i. \end{cases}$$

Assume without loss of generality that $t_1 \leq t_2 \leq \ldots \leq t_m$. Then if $t_i = t_j$, we find that

$$S_{ij} = D_i^T D_j + B_i^T P_0 B_j,$$

with $P_0 = \sum_{t=0}^{\infty} (A^t)^T C^T C A^t$ the observability Gramian. If $t_i < t_j$, then

$$S_{ij} = B_i^T (A^{t_j - t_i - 1})^T C^T D_j + B_i^T (A^{t_j - t_i})^T P_0 B_j.$$

The case $t_i > t_j$ is symmetric.

□

3.4.4 Discussion

In the expression (3.5) another maximization over the inter-event times $t_i - t_j$ still need to be carried out. This optimization depends on the parameters of system G. Therefore, this result could not be directly used in more advanced mechanism optimization schemes, such as the one discussed in the next section. However this could be used to evaluate carefully the amount of noise necessary in an output perturbation mechanism.

For example, the expression (3.5) provides some intuition about the way the MIMO system dynamics influence its sensitivity. Note from the expression of S_{ij}^{τ} in (3.4) that one way of decreasing the sensitivity of G is to increase sufficiently the required time $|t_i - t_j|$ between the events contributed by a single user, in order for $\|A^{|t_i - t_j|}\|$ to be small enough. For example, a lower bound on inter-event times in different streams could be enforced in the adjacency relation, which would weaken the differential privacy guarantee but help design mechanisms with better performance.

3.5 Zero-Forcing MIMO Mechanisms

Using the sensitivity calculations above, we can now design differentially private mechanisms to approximate a given filter F, as discussed in Section 3.3.1. The mechanisms described below generalize to the MIMO case some ideas introduced in [26]. Indeed, the general approximation architecture considered, described on Fig. 3.1, is the same as for the SISO case. On this figure, the system H is of the form $H = FL$, with L a left inverse of the pre-filter G. We call the resulting mechanisms Zero-Forcing Equalization (ZFE) mechanisms. The goal is to design G (and hence, H) so that the Mean-Squared Error (MSE) between y and $\hat{y}$ on Fig. 3.1 is minimized. In order to obtain a differentially private signal v, the Gaussian white noise signal w has its standard deviation proportional to the sensitivity of the filter G. It was shown in [26] that this setup can allow significant performance improvements compared to the output-perturbation mechanism. Note that the

latter can be recovered when $G = F$ and H is the identity.

3.5.1 SIMO system approximation

First, let us assume that F is a SIMO filter, with p outputs. Note that this scenario is considered in [28] (from a very different point of view) for the special case where each row of F is a moving average filter with a different size for the averaging window. Consider a first stage $G(z) = \mathrm{col}(G_1(z), \ldots, G_q(z))$ taking the input signal u and producing q intermediate outputs that must be perturbed. The second stage is taken to be $H = FL$, with $L(z) = [L_1(z), \ldots, L_q(z)]$ a left-inverse of G, i.e., satisfying

$$\sum_{i=1}^{q} L_i(z)G_i(z) = 1.$$

Let us also define the transfer functions M_i, $i = 1, \ldots, q$, such that $M_i(z) = L_i(z^{-1})$, hence $M_i(e^{j\omega}) = L(e^{j\omega})^*$, and thus in particular

$$|M_i(e^{j\omega})|^2 = |L_i(e^{j\omega})|^2, i = 1, \ldots, q, \tag{3.6}$$

$$\text{and } \sum_{i=1}^{q} M_i(e^{j\omega})^* G_i(e^{j\omega}) = 1. \tag{3.7}$$

From Theorem 5, the sensitivity of the first stage for input signals that are adjacent according to (3.1) is $k_1\|G\|_2$. Using Theorem 4, adding a white Gaussian noise w to the output of G with covariance matrix $k_1^2\kappa_{\delta,\epsilon}^2\|G\|_2^2 I_q$ is sufficient to ensure that the signal v on Fig. 3.1 is differentially private. The MSE for this mechanism can be expressed as

$$\begin{aligned}
\xi(G) &= \lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{\infty} \mathbb{E}\left[\|(Fu)_t - (FLGu)_t - (FLw)_t\|_2^2\right] \\
\xi(G) &= \lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{\infty} \mathbb{E}\left[\|(FLw)_t\|_2^2\right] \\
\xi(G) &= k_1^2\kappa_{\delta,\epsilon}^2\|G\|_2^2\|FL\|_2^2.
\end{aligned}$$

We are thus led to consider the minimization of $\|FL\|_2^2\|G\|_2^2$ over the pre-filters G. Recall in the

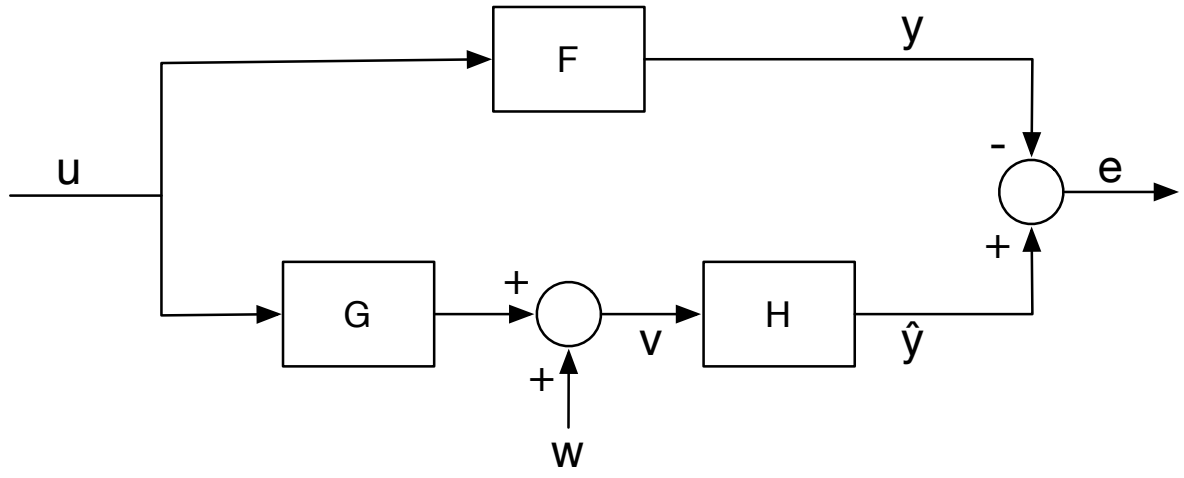


Figure 3.1 Approximation setup for differentially private filtering. w is a noise signal guaranteeing that v is a differentially private signal. The signal $\hat{y}$ is differentially private no matter what the system H is, see [Theorem 3].

following calculation that $|\cdot|$ is also used to denote the Euclidean norm in $\mathbb{C}^d$, for any d. We have

$$\begin{aligned}
&\|FL\|_2^2\|G\|_2^2 \\
=&\frac{1}{2\pi}\int_{-\pi}^{\pi}\mathrm{Tr}(L^*(e^{j\omega})F^*(e^{j\omega})F(e^{j\omega})L(e^{j\omega}))d\omega\times \\
&\frac{1}{2\pi}\int_{-\pi}^{\pi}\mathrm{Tr}(G^*(e^{j\omega})G(e^{j\omega}))d\omega \\
=&\frac{1}{2\pi}\int_{-\pi}^{\pi}|F(e^{j\omega})|^2|L(e^{j\omega})|^2d\omega\times\frac{1}{2\pi}\int_{-\pi}^{\pi}|G(e^{j\omega})|^2d\omega \\
=&\frac{1}{2\pi}\int_{-\pi}^{\pi}|F(e^{j\omega})|^2|M(e^{j\omega})|^2d\omega\times\frac{1}{2\pi}\int_{-\pi}^{\pi}|G(e^{j\omega})|^2d\omega,
\end{aligned}$$

where in the last equality we used (3.6). Now consider on the space of 2π-periodic functions with values in $\mathbb{C}^q$ the inner product

$$\langle f,g\rangle=\frac{1}{2\pi}\int_{-\pi}^{\pi}f(e^{j\omega})^*g(e^{j\omega})d\omega.$$

By the Cauchy-Schwarz inequality for this inner product applied to the functions $\omega\mapsto|F(e^{j\omega})|M(e^{j\omega})$

and $\omega \mapsto G(e^{j\omega})$, we obtain the following bound

$$\|FL\|_2^2\|G\|_2^2 \geq \left(\frac{1}{2\pi}\int_{-\pi}^{\pi}|F(e^{j\omega})|\sum_{i=1}^{q}M_i^*(e^{j\omega})G_i(e^{j\omega})d\omega\right)^2,$$

i.e., using (3.7),

$$\|FL\|_2^2\|G\|_2^2 \geq \left(\frac{1}{2\pi}\int_{-\pi}^{\pi}|F(e^{j\omega})|\,d\omega\right)^2.$$

Moreover, the two sides in the Cauchy-Schwarz inequality are equal, i.e., the bound is attained, if

$$|F(e^{j\omega})|M(e^{j\omega}) = G(e^{j\omega}).$$

Note that this condition does not depend on q. Hence we can simply take $q = 1$, and $L(z) = 1/G(z)$, to get

$$\begin{aligned}&|F(e^{j\omega})|L^*(e^{j\omega}) = G(e^{j\omega})\\ &\text{i.e., } |G(e^{j\omega})|^2 = |F(e^{j\omega})|.\end{aligned} \tag{3.8}$$

Finding G SISO satisfying (3.8) is a spectral factorization problem. We can choose G stable and minimum phase, so that its inverse is also stable. The following theorem summarizes the preceding discussion and generalizes [16, Theorem 8].

Theorem 9. *Let F be a SIMO LTI system with $\|F\|_2 < \infty$. We have, for any LTI system G,*

$$\xi(G) \geq k_1^2\kappa_{\delta,\epsilon}^2\left(\frac{1}{2\pi}\int_{-\pi}^{\pi}|F(e^{j\omega})|d\omega\right)^2, \tag{3.9}$$

where $|F(e^{j\omega})| = \left(\sum_{i=1}^{p}|F_{i1}(e^{j\omega})|^2\right)^{1/2}$. If moreover F satisfies the Paley-Wiener condition $\frac{1}{2\pi}\int_{-\pi}^{\pi}\ln|F(e^{j\omega})|\,d\omega > -\infty$, this lower bound on the mean-squared error of the ZFE mechanism can be attained by some minimum phase SISO system G such that $|G(e^{j\omega})|^2 = |F(e^{j\omega})|$, for almost every $\omega \in [-\pi, \pi)$.

3.5.2 MIMO system approximation

Let us now assume that F has $m > 1$ inputs. We write $F(z) = [F_1(z), \ldots, F_m(z)]$, with F_i a $p \times 1$ transfer matrix. In this case, in view of the complicated expression (3.5) for the sensitivity of a general MIMO filter, we only provide a subpotimal ZFE mechanism, together with a comparison between the performance of our mechanism and the optimal ZFE mechanism. The idea is to restrict our attention to pre-filters G that are $m \times m$ and diagonal, for which the sensitivity is given in

Theorem 6. The problem of optimizing the diagonal pre-filters, using the architecture depicted on Fig. 3.2, can in fact be seen as designing m SIMO mechanisms.

Diagonal Pre-filter Optimization

If G is diagonal, then according to Theorem 6 its squared sensitivity is $(\Delta_2^{m,m}G)^2 = \|KG\|_2^2 = \sum_{i=1}^m \|k_i G_{ii}\|_2^2$, with $K = \mathrm{diag}(k_1, \ldots, k_m)$. Following the same reasoning as in the previous subsection, the MSE for this mechanism can be expressed as

$$\xi(G) = \kappa_{\delta,\epsilon}^2 \|KG\|_2^2 \|FG^{-1}\|_2^2,$$

with $G^{-1}(z) = \mathrm{diag}(G_{11}(z)^{-1}, \ldots, G_{mm}(z)^{-1})$. Now remark that

$$\|FG^{-1}\|_2^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{i=1}^m \frac{|F_i(e^{j\omega})|^2}{|G_{ii}(e^{j\omega})|^2} d\omega.$$

Hence from the Cauchy-Schwarz inequality again, we obtain the lower bound

$$\xi(G) \geq \kappa_{\delta,\epsilon}^2 \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{i=1}^m \frac{|F_i(e^{j\omega})|}{|G_{ii}(e^{j\omega})|} |k_i G_{ii}(e^{j\omega})| d\omega \right)^2$$
$$\xi(G) \geq \kappa_{\delta,\epsilon}^2 \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{i=1}^m k_i |F_i(e^{j\omega})| d\omega \right)^2,$$

and this bound is attained if

$$k_i |G_{ii}(e^{j\omega})| = \frac{|F_i(e^{j\omega})|}{|G_{ii}(e^{j\omega})|},$$
$$\text{i.e. } k_i |G_{ii}(e^{j\omega})|^2 = |F_i(e^{j\omega})|, \ i = 1, \ldots, m.$$

In other words, the best diagonal pre-filter for the MIMO ZFE mechanism can be obtained from m spectral factorizations of the functions $\omega \mapsto \frac{1}{k_i}|F_i(e^{j\omega})|$, $i = 1, \ldots, m$.

Theorem 10. *Let $F = [F_1, \ldots, F_m]$ be a MIMO LTI system with $\|F\|_2 < \infty$. We have, for any diagonal filter $G(z) = diag(G_{11}(z), \ldots, G_{mm}(z))$,*

$$\xi(G) \geq \kappa_{\delta,\epsilon}^2 \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{i=1}^m k_i |F_i(e^{j\omega})|_2 d\omega \right)^2. \tag{3.10}$$

If moreover each F_i satisfies the Paley-Wiener condition $\frac{1}{2\pi}\int_{-\pi}^{\pi} \ln|F_i(e^{j\omega})|\, d\omega > -\infty$, this lower bound on the mean-squared error of the ZFE mechanism can be attained by some minimum phase

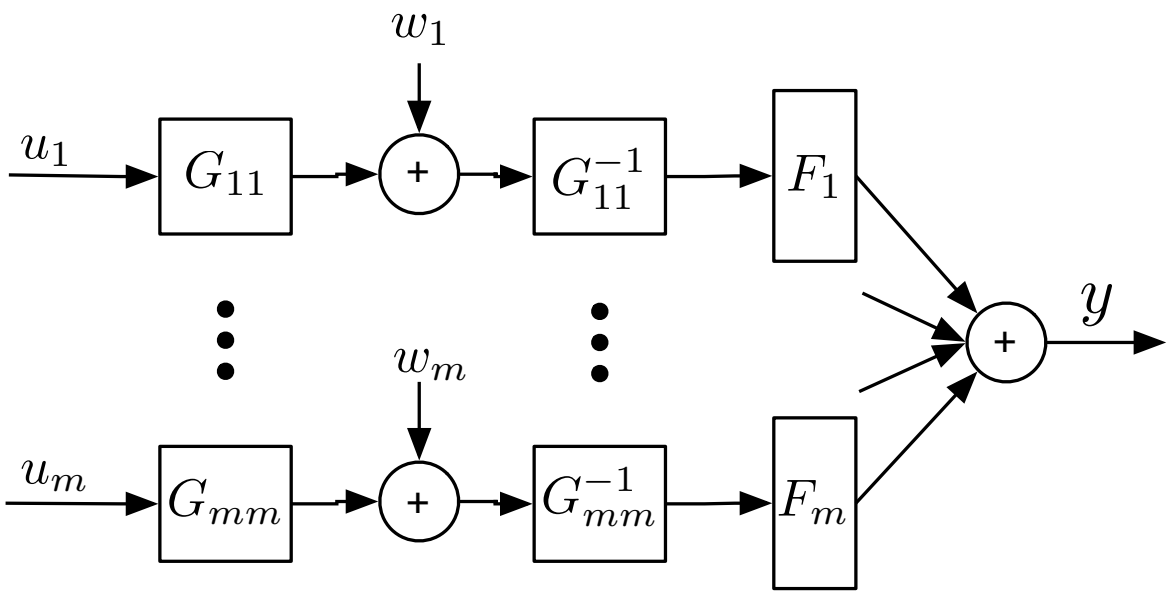


Figure 3.2 (Suboptimal) ZFE mechanism for a MIMO system $Fu = \sum_{i=1}^{m} F_i u_i$, and a diagonal pre-filter $G(z) = \mathrm{diag}(G_{11}(z), \ldots, G_{mm}(z))$. Here $F_i(z)$ is a $p \times 1$ transfer matrix, for $i = 1, \ldots, m$. The signal w is a white Gaussian noise with covariance matrix $\kappa_{\delta,\epsilon}^2 \|KG\|_2^2 I_m$.

systems G_{ii} such that $|G_{ii}(e^{j\omega})|^2 = |F_i(e^{j\omega})|$, for almost every $\omega \in [-\pi, \pi)$.

Comparison with Non-Diagonal Pre-filters

For F a general MIMO system, it is possible that we could achieve better performance with a ZFE mechanism where G is not diagonal, i.e., by combining the inputs before adding the privacy-preserving noise. Here we provide another lower bound on the MSE that one could expect by carrying out this more involved optimization over general pre-filters G rather than just diagonal pre-filters. To simplify the discussion, we assume $k_1 = \ldots = k_m = 1$.

Hence consider a general $m \times m$ pre-filter G with left inverse L. With the lower bound of Theorem 7, designing a ZFE mechanism based on sensitivity as above would require adding a noise with

variance at least $\kappa_{\delta,\epsilon}^2 \|G\|_2^2$. This would lead to an MSE at least equal to $\kappa_{\delta,\epsilon}^2 \|G\|_2^2 \|FL\|_2^2$. Now note that

$$\begin{aligned}
\|FL\|_2^2 &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathrm{Tr}(F(e^{j\omega})L(e^{j\omega})L(e^{j\omega})^* F(e^{j\omega})^*)d\omega \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathrm{Tr}(F(e^{j\omega})^* F(e^{j\omega})L(e^{j\omega})L(e^{j\omega})^*)d\omega \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathrm{Tr}(A(e^{j\omega})^2 L(e^{j\omega})L(e^{j\omega})^*)d\omega \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathrm{Tr}(A(e^{j\omega})L(e^{j\omega})L(e^{j\omega})^* A(e^{j\omega}))d\omega,
\end{aligned}$$

where for all ω, $A(e^{j\omega})$ is the positive-semidefinite square root of $F(e^{j\omega})^* F(e^{j\omega})$, i.e., $A(e^{j\omega})^2 = F(e^{j\omega})^* F(e^{j\omega})$. Then, once again from the Cauchy-Schwarz inequality

$$\begin{aligned}
\|G\|_2^2\|FL\|_2^2 &= \left[\frac{1}{2\pi}\int_{-\pi}^{\pi} \mathrm{Tr}(G(e^{j\omega})^* G(e^{j\omega}))d\omega\right] \\
&\quad \times\left[\frac{1}{2\pi}\int_{-\pi}^{\pi} \mathrm{Tr}(A(e^{j\omega})L(e^{j\omega})L(e^{j\omega})^* A(e^{j\omega}))d\omega\right] \\
\|G\|_2^2\|FL\|_2^2 &\geq \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \mathrm{Tr}(A(e^{j\omega})L(e^{j\omega})G(e^{j\omega}))d\omega\right)^2 \\
\xi(G) &\geq \kappa_{\delta,\epsilon}^2 \left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \|F(e^{j\omega})\|_* d\omega\right)^2, \qquad (3.11)
\end{aligned}$$

where $\|F(e^{j\omega})\|_* = \mathrm{Tr}(A(e^{j\omega}))$ denotes the nuclear norm of the matrix $F(e^{j\omega})$ (sum of singular values). The lower bound (3.11) on the achievable MSE with a general pre-filter in a ZFE mechanism should be compared to the performance (3.10) that we obtained with diagonal pre-filters (with $k_i = 1$ here). Note that these bounds indeed coïncide for $m = 1$.

3.6 Example : Estimation of Building Occupancy

In this section we illustrate some of the ideas discussed above in the context of an application to estimating and forecasting occupancy in an office building equipped with motion detection sensors. As mentioned in Section 3.3.2, such an application raises privacy concerns related to the possibility that some occupants could be tracked individually from the published information, correlated possibly with public information such as the location of their office. The dataset used here comes from a sensor network experiment carried out in the Mitsubishi Electric Research Laboratories (MERL) and described in [2] and on Fig. 3.3.

The original dataset contains the traces of more than 200 sensors spread over two floors of a building, where each sensor recorded with millisecond accuracy over several months the exact times

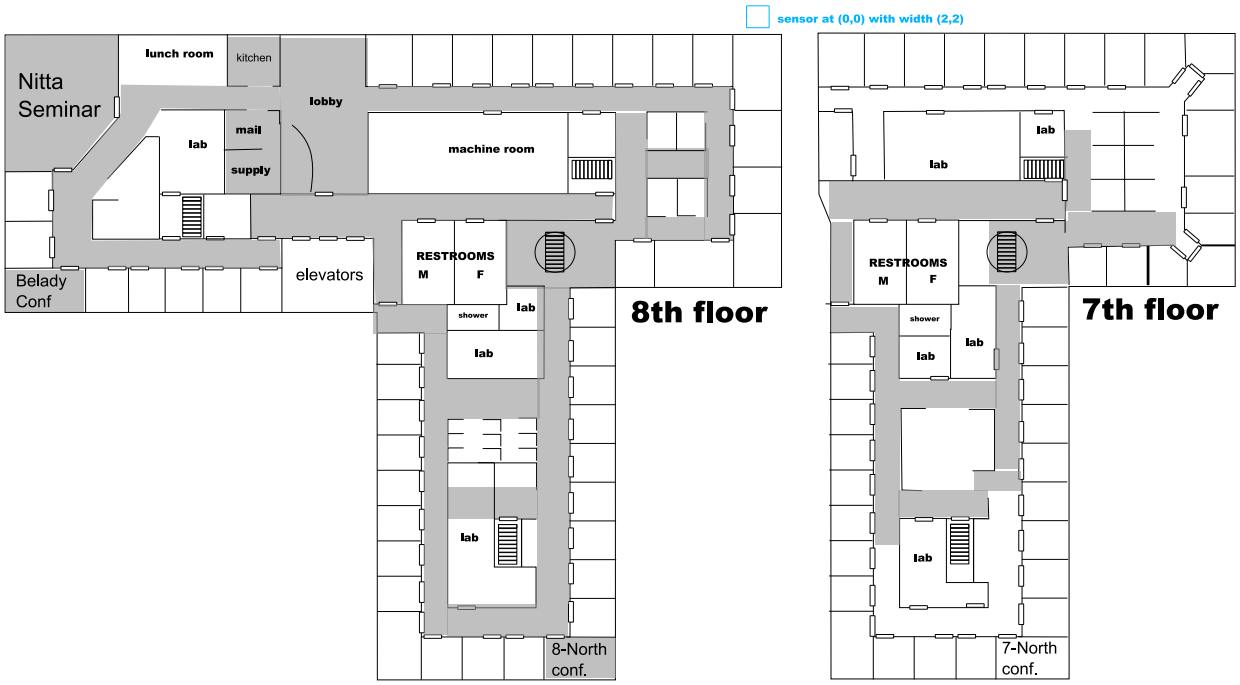


Figure 3.3 Floorplan of the part of the MERL building used for the sensor network experiment in [2]. The shaded areas are hallways, lobbies and meeting rooms equipped with more than 200 motion detectors, placed a few meters apart and recording events roughly every second.

at which they detected some motion. For illustration purposes we subsampled the dataset in space and time, summing all the events recorded by several sufficiently close sensors over 5 minute intervals. We formed in this way 10 input signals u_i, $i = 1, \ldots, 10$, corresponding to 10 spatial zones (each zone covered by a group of several sensors), with a discrete-time period corresponding to 5 minutes, and $u_{i,t}$ being the number of events detected by all the sensors in group i during period t. If us assume say that during a given discrete-time period, a single individual can activate at most 2 sensors in any group, then $k_i = 2$ for $1 \leq i \leq 25$. If moreover we assume that any individual travels through at most 5 zones, then we could add a constraint $\sum_{i=1}^{10} k_i \leq 10$. Finally we need to assume that a single individual only activates a group of sensor once over the time interval for which we wish to provide differential privacy. Section 3.3.2 discussed how to relax this requirement by splitting the input data.

As an illustrative example, we could be interested in a system computing simultaneously and in real-time the following three outputs :

– The sum of the moving averages over the past 30 min for zones 1 to 4.
– The sum of the moving averages over the past 1 h for zones 3 to 7.
– A forecast (prediction) of the total number of events detected in the next 20 minutes in all zones, provided by an ARMAX model [29] (with 10 inputs and one output) calibrated using one part of the dataset.

Hence our desired filter has the structure

$$F = \begin{bmatrix} * & * & * & * & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & * & * & * & * & * & 0 & 0 & 0 \\ * & * & * & * & * & * & * & * & * & * \end{bmatrix}, \tag{3.12}$$

where $*$ denotes a non-zero transfer function. Fig. 3.4 shows a sample path over a 25 h period of the 3rd output for a predictive ARMAX model that we designed, as well as a differentially-private version obtained from the diagonal pre-filter optimization procedure of Section 3.5.2 applied to the whole filter F. Notice on the figure that the approach used here relying on the notion of sensitivity requires a noise level independent of the size of the desired output signal, hence low signal values tend to be easily buried in the noise.

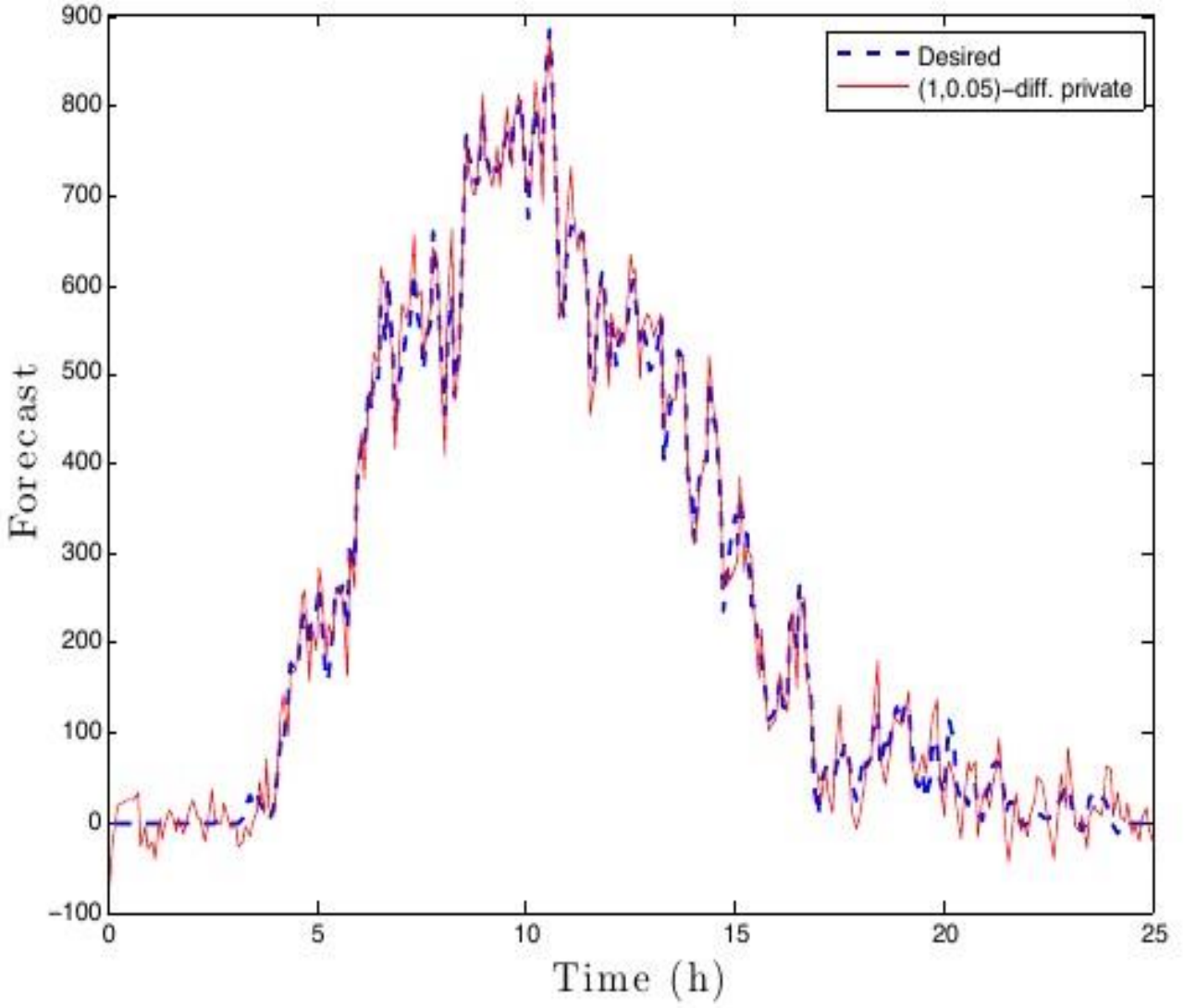


Figure 3.4 Real-time forecast of the total number of events that will be detected in the next 20 minutes in the whole building. We show the output of an ARMAX model calibrated on the dataset, which is the 3rd output of the filter F in (3.12). We also show a $(1, 0.05)$-differentially private output.

CHAPITRE 4 DIFFERENTIALLY PRIVATE TRAFFIC STATE ESTIMATION

4.1 Introduction

Managing traffic by means of road traffic information systems could resolve one of the major concerns in urban areas, namely, *traffic congestion* . This phenomenon is responsible for the heavy costs linked to lost time, fuel consumption and increased pollution. One of the most important strategy implemented in traffic information systems to avoid unnecessary costs of developing new expensive infrastructures is to apply real-time control policies for demand management. These strategies can be enabled by proliferation of swarms of sensors a long the highways to provide an accurate picture of the traffic state over time. According to the vast literature on traffic estimation, such traffic state estimators mainly use the traffic data obtained from inductive loop detectors embedded in the highway. However, the detailed data from these detectors could be linked to some external information in order to establish presence or absence of individuals contributing this data. The main privacy concern with designing the traffic systems services is the possibility of reconstructing the trajectory of a given person by linking the vast sources of side information to the aggregated data, e.g, real-time traffic density maps, published by the traffic estimators. This chapter aims at designing a traffic state estimator with a privacy-preserving scheme relying on the notion of differential privacy. The scheme presented in this thesis relies on both the microscopic data originating from the participants and a macroscopic model of the aggregate dynamics, namely the *cell transmission model* (CTM) [30].

Briefly, CTM anticipates macroscopic traffic behavior on a given highway in time by evaluating traffic parameters (flow and density) at a finite number of intermediate points at different time steps [30]. This procedure is done by dividing the highway into homogeneous sections (cells) and evaluating the discrete model of traffic flow dynamics at each cell continuously. The cell transmission model together with the measurements data obtained from single loop detectors are then assimilated through an Extended Kalman Filter (EKF), which provides a traffic density map of the spatial interval of interest [30]. The rest of this chapter is organized as follows. Section 4.2 presents a brief background on privacy preserving traffic state estimators and the main motivations behind our work. Sections 4.3 and 4.4 present some background on the traffic flow model and the measurement models for the data originating from single-loop detectors which are necessary to develop a model-based traffic state estimator. Section 4.5 introduces our method for traffic state estimation, built around an extended Kalman filter and a traffic mode estimator. Section 4.6 describes our private mode and density measurement model. Finally, the overall architecture of our private traffic estimator and the constructed private density map are presented in Section 4.7 . We also compare

our results to those presented in [3].

4.2 Background

The problem of designing a differentially private traffic density estimator is initially addressed in [3]. The main drawback of this scheme is to leak the participants information as the number of the sensors (or correspondingly the length of the spatial interval of interest) increases. The mechanism proposed in [3] provides differential privacy guarantee which is inversely related to the number of the sensors reporting the traffic data. For example, for a road equipped with m single loop detectors and any choice of ϵ and δ, the architecture would present a traffic density map with $(M + \epsilon, \delta)$-differential privacy guarantee. In particular this chapter is inspired by [3] and aims at designing a privacy preserving traffic estimator such that the differential privacy guarantee provided to the users data becomes independent of the number of the sensors from which measurements are obtained.

Our result entails a strict improvement over the state-of-the-art: we propose an architecture that could monitor the traffic status continuously with $(2\epsilon, 2\delta)$-differential privacy guarantee where ϵ and δ are chosen by the designer. Furthermore, our mechanism design allows that a big part of the dataset remain non-randomized. This distinguishing feature of the mechanism helps reduce the degradation in estimation performance.

4.3 Traffic Flow Dynamics

The unidirectional traffic along a single road section, with position denoted x and for varying number of lanes $\lambda(x)$ can be mapped based on the traffic flow dynamics [31]

$$q(x,t) = \rho(x,t)v(x,t)$$

where ρ is the vehicle density (say, in vehicles per mile) over all lanes, q is the traffic flow over all lanes, and v is the traffic velocity. We then consider the simplest situation, assuming a homogeneous road section [31]. The continuity equation reads

$$\frac{\partial \rho}{\partial t} + \frac{\partial(\rho v)}{\partial x} = 0 \tag{4.1}$$

We consider instead its discrete version, assuming the road section has been divided into cells of length Δx_i and a time step of Δt [31]. The density in cell i over all lanes follows the recursion

$$\rho^i(t + \Delta t) = \rho^i(t) + \frac{\Delta t}{\Delta x^i}(F_{tot}(\rho^{i-1}(t), \rho^i(t)) - F_{tot}(\rho^i(t), \rho^{i+1}(t)) \tag{4.2}$$

where $F_{tot}(\rho^{i-1}(t), \rho^{i}(t))$ is the total so-called numerical flux enters cell i [30] (i.e., through the interface $i-1 \rightarrow i$) during period Δt, and $F_{tot}(\rho^{i}(t), \rho^{i+1}(t))$ is the total numerical flux out of the cell i (i.e., through the interface $i \rightarrow i+1$). Note here that the numerical flux $F_{tot}(\rho^{i}(t), \rho^{i+1}(t))$ is different in general from the total flow $q(x_{i|i+1}, t)$, where $x_{i|i+1}$ denotes the location of the interface between cells i and $i+1$. More details are provided below.

To complete the model, we need to make an hypothesis on the relationship between two quantities, e.g., between velocity and density, or between flow and density. In order to fulfil this objective, we first introduce lane-averaged, also called effective, quantities, i.e., lane-averaged traffic density $\rho(x,t)$ (say, in vehicles per mile per lane), lane-averaged traffic speed $v(x,t)$, and lane-averaged traffic flow $q(x,t) = \rho(x,t)v(x,t)$ [31, Chapter 7]. Denoting by $\rho_j(x,t)$, $q_j(x,t)$ and $v_j(x,t)$ the density, speed and flow in lane j at a position x, we have the relations

$$\rho(x,t) = \frac{\sum_{j=1}^{\lambda(x)} \rho_j(x,t)}{\lambda(x)},\ q(x,t) = \frac{\sum_{j=1}^{\lambda(x)} q_j(x,t)}{\lambda(x)},\ v(x,t) = \frac{\sum_{j=1}^{\lambda(x)} v_j(x,t)}{\lambda(x)}$$

We then adopt the simplest approach based on the first order model which considers a static relationship $q(\rho)$, also called a fundamental diagram. In first-order models, proposed initially by Lighthill and Whitham [32] and independently by Richards [33] (LWR models), the effective density is a fundamental quantity and a sufficient description of the local traffic state, since the effective speed and thus also the effective flow are assumed to be known static functions of density [3]. Note that LWR models assume that the traffic flow is always in local equilibrium with respect to the density, and leads to the formation of physically impossible phenomena such as shock waves. Here, we work for concreteness with triangular fundamental diagrams, which are arguably the most popular in practice. Next we describe the resulting LWR model which is also called the Cell-Transmission Model (CTM) [30].

4.3.1 Cell Transmission Model

The simplest LWR model, called the cell-transmission model, uses a triangular fundamental diagram

$$q(\rho) = \begin{cases} v_f \rho & if \quad \rho \leq \rho_c \\ w(\rho_{max} - \rho) & if \quad \rho_c \leq \rho \leq \rho_{max} \end{cases} \tag{4.3}$$

Here v_f is the velocity of free traffic (say $110 km/h$ for a highway), $\rho_{max}$is the maximum density on this road segment (say 120 vehicles/lane/km for a highway) and ρ_c is the critical density at which the maximum flow $q_{max} = v_f \rho_c$ is attained. Note that w is the velocity of the waves of density variations in congested traffic (which propagate backwards). Fig. 4.1 illustrates these definitions[31].

Dividing the road into I cells numbered $1, ..., I$, the discrete-time lane-averaged conservation law

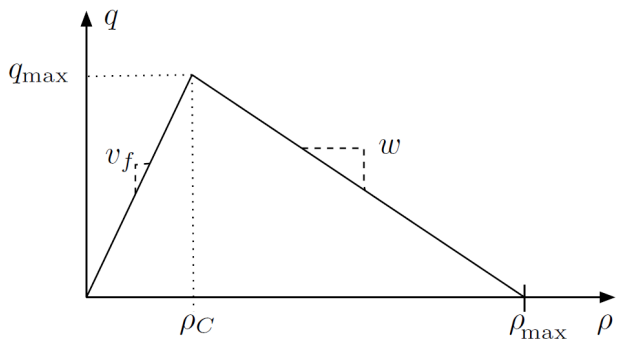


Figure 4.1 Triangular fundamental diagram and associated parameters

for vehicles corresponding to the solution $q(\rho)$ of equation (4.1) is

$$\rho_{k+1}^i = \rho_k^i + \frac{\Delta t}{\Delta x_i}\left(\frac{\lambda^{i-1}}{\lambda^i}F(\rho_k^{i-1},\rho_k^i) - F(\rho_k^i,\rho_k^{i+1})\right), \text{ for } i = 1,...,I$$

where ρ_k^i is the lane-averaged vehicle density in cell i at period k, i.e., during the time interval $[k\Delta t,(k+1)\Delta t]$, and $F(\rho_k^i,\rho_k^{i+1})$ is the lane-averaged numerical flux out of the cell i (i.e.,through the interface $i \to i+1$) during period k. We also define λ^i to be the number of lanes at the interface $i \to i+1$. Any location where the number of lanes changes is always assumed to fall inside a cell. This leads to a dynamical system with non-linear (piecewise linear) dynamics. At the end of the road for which we are estimating the traffic, we add two ghost cells 0 and $I+1$ to enforce the boundary conditions. We assume that we have loop detectors at the exit of cell 0 and at the entrance of cell $I+1$, in order to enforce the boundary conditions[30]. Then, the following stochastic state-space model of the density dynamics on the road is obtained

$$\rho_{k+1}^i = \rho_k^i + \frac{\Delta t}{\Delta x_i}\left(\frac{\lambda^{i-1}}{\lambda^i}F(\rho_k^{i-1},\rho_k^i) - F(\rho_k^i,\rho_k^{i+1})\right) + \gamma_k^i, \text{ for } i = 1,...,I \tag{4.4}$$

Here γ_k^i is a Gaussian random variable, whose variance can be tuned later on in the design of the state estimator, based on the relative confidence we place in the model or the observations. The dynamics of the ghost cells are also

$$\rho_{k+1}^0 = \rho_k^0 + \gamma_k^0, \quad \rho_{k+1}^{I+1} = \rho_k^{I+1} + \gamma_k^{I+1} \tag{4.5}$$

Finally, for the triangular fundamental (4.3), the standard numerical method, the Godunov method, corresponds to using the following numerical flux in (4.4):

$$F(\rho_k^i,\rho_k^{i+1}) = min(\rho_k^i v_f, \rho_c v_f, w(\rho_{max} - \rho_k^{i+1})) \tag{4.6}$$

4.3.2 Extended Kalman Filter

As noted in introduction, the stochastic state-space model 4.4 together with the measurements reported by single-loop detectors (Section 4.4), could be assimilated in an extended Kalman filter (EKF) to construct our traffic density map. Here, we present a very short introduction to the EKF. Consider the following non-linear stochastic state space system.

$$x_{k+1} = F(x_k) + \omega_k \qquad k \in \mathbb{Z}_+ \tag{4.7}$$

$$y_k = H(x_k) + \nu_k \tag{4.8}$$

where $x_0 \sim \mathcal{N}(0, \Sigma)$ and that x_0 is independent of the system disturbance process ω and the observation noise process ν. For these two we also assume

$$\begin{bmatrix} \omega \\ \nu \end{bmatrix} \sim \mathcal{N}\left[\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix}\right]$$

A popular approach to the stochastic state estimation for system (4.7) is the extended Kalman filter (see e.g. Chapter 13, [34]). Subject to the assumption that F and H have continuous first order partial derivatives, one may recursively employ the Taylor series expansion of F and H to obtain linear approximations of the system dynamics process and observations process in the neighbourhood of the time varying trajectory $x_k, k \in \mathbb{Z}_+$. Henceforth, this assumption will be adopted without further comment. Carrying out this first order approximation for $F(x_k)$ the estimated state $\hat{x}_k$ could be obtained with the following conditioning and prediction steps.
Conditioning step:

$$\hat{x}_k = x_{k|k-1} + V_k H_k^T \left[H_k V_k H_k^T + R\right]^{-1} (y_k - H(x_{k|k-1}))$$

Prediction step:

$$V_{k+1} = F_k V_k F_k^T - F_k^T H_k^T \left[H_k V_k H_k^T + R\right]^{-1} H_k F_k + Q$$

where

$$x_{k+1|k} = F(\hat{x}_k), \qquad V_0 = \Sigma, \qquad F_k = \left[\frac{\partial F(x)}{\partial x}\right]|_{x=\hat{x}_k}, \qquad H_k = \left[\frac{\partial H(x)}{\partial x}\right]|_{x=x_{k|k-1}}$$

4.4 Single Loop Detector Measurement Model

The datasets provided by the flow sensors consist of sequences of counts $c^i_{j,k}$,and occupancies $o^i_{j,k}$ for $k \geq 0, 1 \leq i \leq S$,$0 \leq j \leq \lambda^i$. Here k represents the period ($T = 30s$), S is the number of single loop detectors reporting the records, and j determines the lane number. The occupancy $0 \leq o^i_{j,k} \leq 1$ is a unit less number representing the percentage of period k for which a vehicle was passing in front of sensor i. Since single loop detectors cannot measure traffic density or velocity at their location, their measurements must be used to obtain an estimation of these quantities. For let's say a one-lane road equipped with a number of single loop detectors, these estimations are

$$v_j(t) \approx g\frac{c_j(t)}{o_j(t)T}, \;\; q_j(t) \approx \frac{c_j(t)}{T}, \;\; \rho_j(t) \approx \frac{o_j(t)}{g} \tag{4.9}$$

where T is the time period of the sensor and g is the so-called g-factor, which is the average effective vehicle length at the sensor location that can vary with time. Similar to [3], to get a more robust approximation of density, we first form the approximate flows based on the counts data. That is, the flow ϕ^i_k around the sensor placed at the interface $i \rightarrow i+1$ for cells i and $i+1$ is defined by the non-linear measurement model

$$\phi^i_k := \frac{1}{\lambda^i T}\sum_{j=1}^{\lambda^i} c^i_{j,k} = F(\rho^i_k, \rho^{i+1}_k) + \nu_k \tag{4.10}$$

where ν_k is a Gaussian random variable describing the measurement errors. The density pseudo-measurement model is then defined as

$$z^i_k = z^{i+1}_k = \begin{cases} \frac{\phi^i_k}{v_f} & if \quad m^i_k = F \\ \rho_{max} - \frac{\phi^i_k}{w} & if \quad m^i_k = C \end{cases} \tag{4.11}$$

where m^i_k is the traffic mode for the interface, either free (F) or congested (C) corresponding to $\rho \leq \rho_c$ and $\rho \geq \rho_c$. In fact, this model is obtained by inverting our triangular fundamental diagram presented in (4.3). The observation signal z^i_k is related simply to the density of the flow as

$$z^i_k = z^{i+1}_k = \rho^i_k + \eta^i_k = \rho^{i+1}_k + \eta^{i+1}_k \tag{4.12}$$

where η^i_k, η^{i+1}_k are assumed to be Gaussian random variables. However, this model requires determining the exact mode of the traffic flow. The strategy proposed in [3] to find the mode of the traffic is to use the reported occupancy measurements to estimate the mode of the traffic either fluid or congested, which corresponds to $\frac{o_j}{g} \leq \rho_c$ or $\frac{o_j}{g} > \rho_c$ respectively.

However, these traffic mode measurements can result in frequent mode estimation errors due simply to an inaccurate estimation of g-factor. With 18-foot long autos and 60-foot long trucks, the g-factor parameter is expected to range from 18 feet for inner, auto-only lanes to as much as 60 feet in the early morning for outer lanes over fluid highways with heavy truck traffic [35]. Note also that these mode measurements appear to be difficult to handle from a differential privacy point of view, because the occupancy time due to a single vehicle, equal to $\frac{l_v}{Tv_v}$, with l_v its length and v_v its speed, can vary widely depending on its speed. As a result, the sensitivity of these occupancy measurements is high and the standard Gaussian perturbation mechanism leads to more unreliable measurements, especially at low density [3]. We now present our mode measurement model, which takes both the occupancy and the count measurements into account to obtain a more reliable estimation of the traffic modes.

4.5 Non-Private Mode and Density Measurements

According to (4.11), to each flow measurement $0 \leq \phi_k^i < q_{max}$(4.10) correspond two possible densities on the fundamental diagram. Based on (4.9), we can also form the lane-average contribution to the density via occupancy measurements as

$$y_k^i = \frac{1}{g\lambda^i} \sum_{j=1}^{\lambda^i} o_{j,k}^i \tag{4.13}$$

The traffic mode pseudo-measurements can be then obtained as

$$M_k^i = M_k^{i+1} = \arg\min_{m_k^i} \left(\left| z_k^i(m_k^i) - y_k^i \right| \right) \tag{4.14}$$

Actually, model (4.14) estimates the mode either free (F) or congested (C) based on which sub function in the hybrid function $z_k^i(m_k^i)$ (4.11) is closer to the occupancy contribution to the density y_k^i. However, this model requires estimating the g-factor precisely to guarantee that the minimum term in expression (4.14) is correctly chosen. This is problematic, because the g-factor can vary with time and generally is not easy to estimate. To handle this problem, we now assume that the g-factor is a constant and stay the same over time, namely 20 feet. We then bound the allowed deviation between the density pseudo-measurement z_k^i (4.3) and the occupancy contribution to density y_k^i (4.13) . That is

$$\text{for some constant } g,\ \exists \zeta(g) > 0 \ s.t.\ \left| \log[z_k^i] - \log[y_k^i] \right| \leq \zeta(g) \qquad \forall i,k \tag{4.15}$$

Note that, truncation (4.15) transforms the variation in g-factor into the upper-bound error between z_k^i and y_k^i, that is

$$\left|\log \frac{z_k^i}{y_k^i}\right| \leq \zeta(g) \equiv \frac{1}{\mathbf{g}\mathbf{e}^{\zeta(\mathbf{g})}\lambda^i} \sum_{j=1}^{\lambda^i} o_{j,k}^i \leq z_k^i \leq \frac{1}{\frac{\mathbf{g}}{\mathbf{e}^{\zeta(\mathbf{g})}}\lambda^i} \sum_{j=1}^{\lambda^i} o_{j,k}^i \tag{4.16}$$

For assumed values of g and $\zeta(g)$, we then define the sets

$$T_F = \left\{ \left(\phi_k^i, y_k^i\right) \ : \ \left|\log\left[\frac{\phi_k^i}{v_f}\right] - \log\left[y_k^i\right]\right| \leq \zeta(g) \ \ \forall i,k \right\} \tag{4.17}$$

$$T_C = \left\{ \left(\phi_k^i, y_k^i\right) \ : \ \left|\log\left[\rho_{max} - \frac{\phi_k^i}{w}\right] - \log\left[y_k^i\right]\right| \leq \zeta(g) \ \ \forall i,k \right\} \tag{4.18}$$

corresponding to the flow ϕ_k^i satisfying our truncation in free mode (F) or congested mode (C) respectively.

Lemma 11. *For any flow ϕ_k^i, defined in (4.10), we have*

$$\mathbf{1}_{T_F}\left((\phi_k^i, y_k^i)\right)\mathbf{1}_{T_C}\left((\phi_k^i, y_k^i)\right) = 1 \quad \textit{iff} \quad \phi_k^i \in \left[\frac{w v_f \rho_{max}}{w e^{2\zeta(g)} + v_f}, \frac{w e^{2\zeta(g)} v_f \rho_{max}}{w + e^{2\zeta(g)} v_f}\right]. \tag{4.19}$$

Proof. We have

$$e^{-\zeta(g)}\frac{\phi_k^i}{v_f} \leq y_k^i \leq e^{\zeta(g)}\frac{\phi_k^i}{v_f}, \quad \text{for all } (\phi_k^i, y_k^i) \in T_F \tag{4.20}$$

$$e^{-\zeta(g)}(\rho_{max} - \frac{\phi_k^i}{w}) \leq y_k^i \leq e^{\zeta(g)}(\rho_{max} - \frac{\phi_k^i}{w}), \quad \text{for all } (\phi_k^i, y_k^i) \in T_C \tag{4.21}$$

In view of (4.20),(4.21), we have

$$T_F \cap T_C = \left\{ (\phi_k^i, y_k^i) \ : \ e^{-\zeta(g)}\frac{\phi_k^i}{v_f} \leq e^{\zeta(g)}(\rho_{max} - \frac{\phi_k^i}{w}) \ \text{ and } \ e^{-\zeta(g)}(\rho_{max} - \frac{\phi_k^i}{w}) \leq e^{\zeta(g)}\frac{\phi_k^i}{v_f} \right\}$$

and by solving these two inequalities for ϕ_k^i

$$\mathbf{1}_{T_F}\left((\phi_k^i, y_k^i)\right)\mathbf{1}_{T_C}\left((\phi_k^i, y_k^i)\right) = 1 \quad \text{iff} \quad \phi_k^i \in \left[\frac{w v_f \rho_{max}}{w e^{2\zeta(g)} + v_f}, \frac{w e^{2\zeta(g)} v_f \rho_{max}}{w + e^{2\zeta(g)} v_f}\right].$$

□

Defining $\bar{T}_C$ and $\bar{T}_F$, the complement sets of T_C and T_F respectively, we could form a traffic mode

measurement model as

$$M_k^i = \begin{cases} F & if \quad \mathbf{1}_{T_F - T_C}\left((\phi_k^i, y_k^i)\right) = 1: \text{ Safe zone Free mode} \\ C & if \quad \mathbf{1}_{T_C - T_F}\left((\phi_k^i, y_k^i)\right) = 1: \text{ Safe zone Congested mode} \\ M_{k-r}^i & if \quad \left[\prod_{s=0}^{r-1} \mathbf{1}_{T_C \cap T_F}\left((\phi_{k-s}^i, y_{k-s}^i)\right)\right] \mathbf{1}_{\bar{T}_C \cup \bar{T}_F}\left((\phi_{k-r}^i, y_{k-r}^i)\right) = 1,\ r > 0: \text{ Sensitive zone} \end{cases} \tag{4.22}$$

The mode measurement model (4.22) determines the mode of the traffic with respect to truncation (4.15), that is, the current mode is either free (F) or congested (C), if the current flow satisfies (4.22) only in free mode or only in congested mode, respectively. The third case corresponds to the flow $\phi_k^i \in \left[\frac{w v_f \rho_{max}}{w e^{2\zeta(g)} + v_f}, \frac{w e^{2\zeta(g)} v_f \rho_{max}}{w + e^{2\zeta(g)} v_f}\right]$ where the truncation is respected in both traffic modes. For this case, the strategy adopted here is to take the mode of the last flow ϕ_{k-r}^i, that is inside one of the two safe zone in (4.22). For illustrative purposes, the region corresponding to the flows satisfying the truncation in one mode on a triangular fundamental diagram is re-presented with a green line on Fig. 4.2 . We call the region corresponding to these flows, the **Safe** zone. We also introduce a **Sensitive** zone (red line) which illustrates the flows that our model is disable to determine its mode directly. The parameters related to the fundamental diagram and the model that are used in Fig. 4.2 are $v_f = 65$ mph, $w = 11.6$ mph, $\rho_{max} = 193$ vehicles/mile/lane, $g = 20$ feet and $\zeta(g) = 0.51$ which is a unit-less parameter. Note that based on (4.16), the choice of $\zeta(g) = 0.51$ will consider g-factor variations between 12 feet and 33.3 feet. Note also that, our mode measurement model (4.22) basically estimates the mode of the traffic with respect to the flows falling inside the safe zone. Although this strategy could reduce the accuracy of the mode measurement, especially for flows in sensitive zone, we show that such a model can be used efficiently in designing a differentially private traffic estimator in Section. 4.6 .

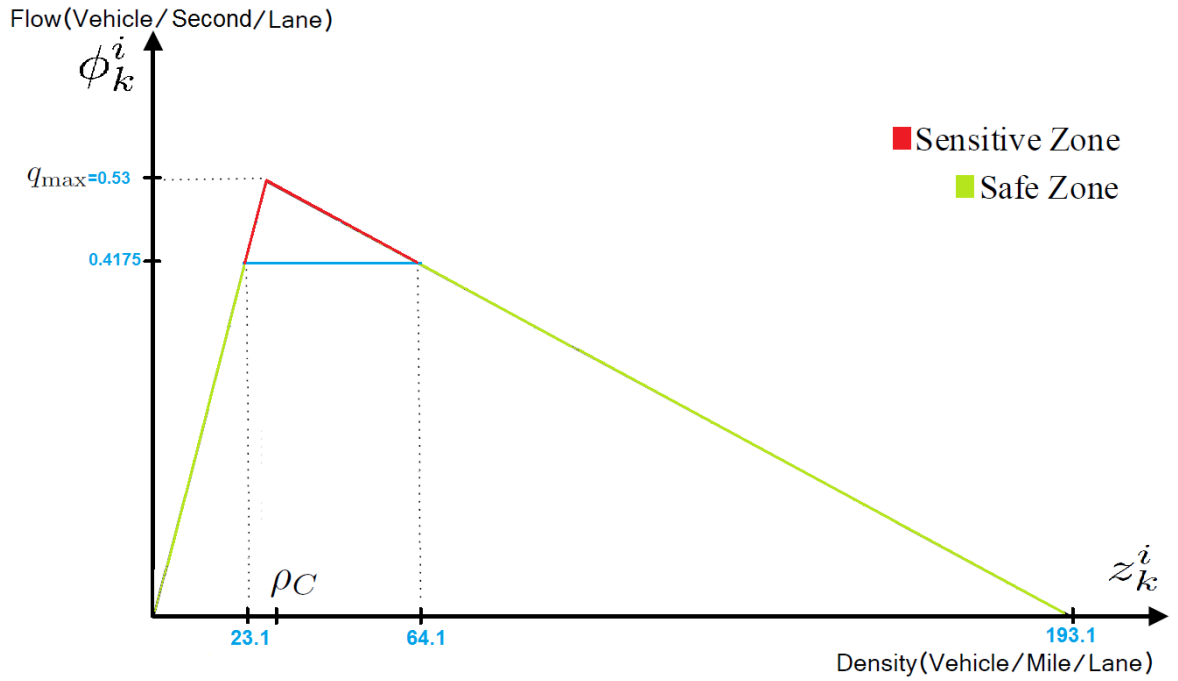


Figure 4.2 Safe zone and Sensitive zone on Triangular Fundamental Diagram for $g = 20$ feet and $\zeta(g) = 0.51$. The safe zone is quite large for simulation purposes.

Similar to [3], to obtain a more physically meaningful mode, the mode measurements will then be filtered through an additional hidden-Markov model (HMM), described as follows. For the state trajectory M_k^i defined in (4.22), the actual mode estimate used to invert the fundamental diagram is introduced by the new state trajectory $\{s_k^i\}_{k\geq 0}$ with $s_k^i \in \{C, F\}$. The dynamics of s_k^i are simply described by a Markov chain with a single parameter $\pi_1 = \mathbb{P}(s_{k+1}^i \neq s_k^i)$, describing the probability with which the mode changes from fluid to congested at that location. This parameter could be estimated from historical data. Finally, we introduce a last parameter $\pi_2 = \mathbb{P}(m_k^i = s_k^i)$, which reflects the confidence we have in the output of the our model. Accordingly, the confidence probability parameter in HMM is then set with respect to the value of flow data, i.e, this parameter for the flows located in sensitive zone is much lower than the ones in safe zone. For the non-private estimation, we could also define the confidence probability as $\pi_2 = \mathbb{P}(m_k^i = s_k^i | q_k^i)$ which is a helpful model for the sensitive zone flows, indeed, even if M_{k-r}^i addresses a wrong mode, the confidence probability will be set according to the occupancy contribution to the density (4.13) and the (HMM) could correct the error. The process of providing density measurements from occupancies and counts data is described in Algorithm. 1.

1- Calculate flow measurements $\phi_k^i = \frac{1}{T\lambda^i} \sum_{j=1}^{\lambda^i} c_{j,k}^i$
2- based on historical data choose a base g-factor, e.g, 20 feet, and an upper-bound error $\zeta(g)$.
3- Specify the corresponding Safe zone and Sensitive zone based on Theorem. 11.
4- $m_k^i = F/C$ based on the mode measurement model (4.22).
5- Filter m_k^i thorough the HMM filter to obtain the actual mode s_k^i used to invert the fundamental diagram.
6-

$$z_k^i = z_k^{i+1} = \begin{cases} \frac{\phi_k^i}{v_f} & if \quad s_k^i = F \\ \rho_{max} - \frac{\phi_k^i}{w} & if \quad s_k^i = C \end{cases} \tag{4.23}$$

Algorithm 1: Non-private density measurement

To illustrate our approach, we estimate the traffic state from induction loop data available as part of the Mobile Century experiment dataset [36]. This data consists of counts and occupancy measurements from single loop detectors, for each lane of Interstate 880 (Northbound) in California between post-mile 16.5 and 27.7, i.e., along an approximately 11 mile road segment. The density measurements based on Algorithm 1 is then assimilated in an Extended Kalman Filter (EKF) to construct our non-private density map.

Fig. 4.3 presents the performance of our non-private estimator. The results are greatly similar to the map presented in Fig. 4.4 which is the non-private map introduced by [3]. This similarity proves the reliability of our mode measurement model. The two maps have some discrepancies mainly in the areas where the traffic is about to switch between the modes. It can be claimed that Fig. 4.3 presents a more reliable picture of the traffic density since our mode measurement model considers the probable variations in g-factor over time. Next, we show that this model can be used efficiently in a differentially private scheme.

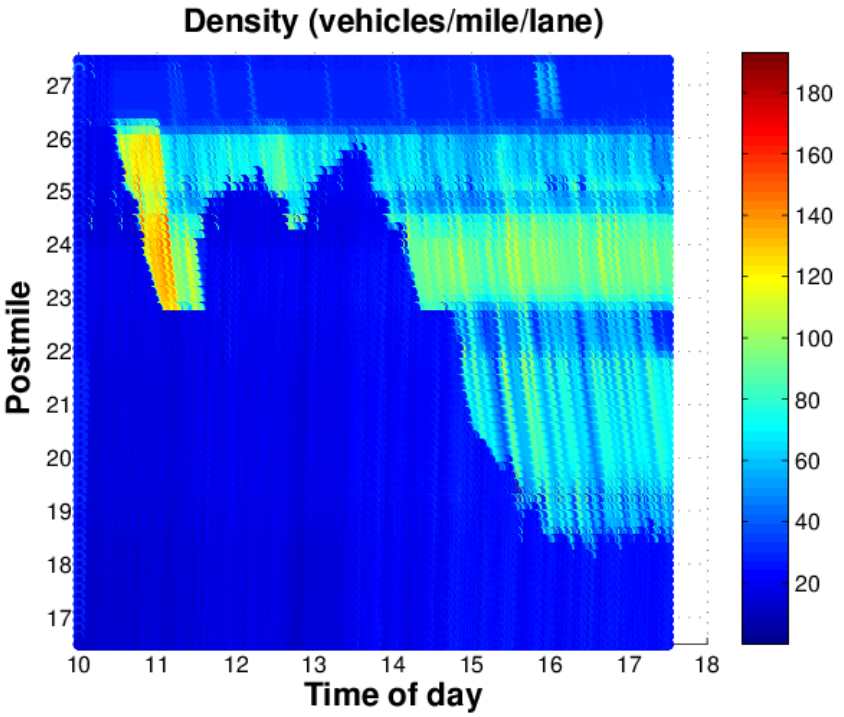


Figure 4.3 Real-time density map reconstruction with a non-private extended Kalman filter based on Algorithm. 1

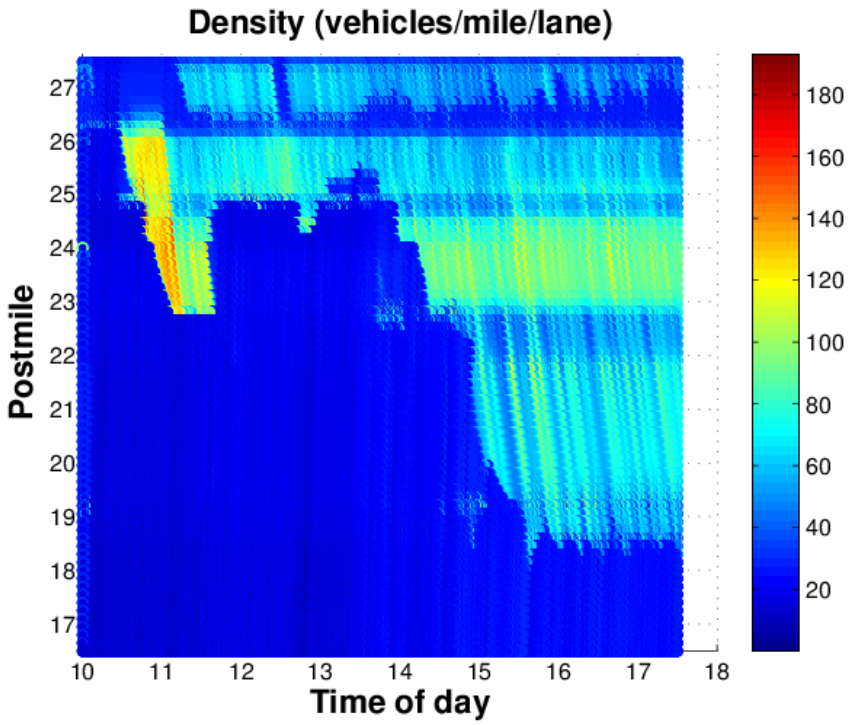


Figure 4.4 Real-time density map reconstruction with a non-private extended Kalman filter presented in [3]

4.6 Differentially Private Mode and Density Measurements

The measurements obtained from the single loop detectors, i.e, counts $c^i_{j,k}$, and occupancies $o^i_{j,k}$, cannot be directly used in any traffic estimator architecture, because they could reveal private information about people who contribute these measurements. In this section, we present differentially private algorithms that output the private flow and traffic mode measurements. These sanitized pseudo-measurements are then used in providing differentially private density pseudo-measurements which is a sufficient observation signal to construct our differentially private density map. The mechanism processing the counts data in order to provide differentially private flows is a simple Gaussian mechanism (3) which was first presented in [3].
In order to provide the private traffic mode measurements, we introduce a new mechanism for sanitizing data sequences which is mainly based on Algorithm 1.
Next, we first review the Gaussian mechanism providing the private flow measurements. Then, the mechanism that provides the mode pseudo-measurements is presented.

4.6.1 Flow Measurements

Similar to [3], the mechanism processing the counts data in order to provide private flow measurements can be a Gaussian mechanism (3). We can consider the following adjacency relation for the counts datasets of N user trajectories $C := \{c^i_{j,k} \ : \ k \geq 0, 1 \leq i \leq S, 1 \leq j \leq \lambda^i\}$

$$\begin{aligned} &\text{for all } c,\ \tilde{c} \in C: \quad Adj(c,\tilde{c}) \ \text{ iff } \ \forall k \geq 0,\ \ \forall i \in [1,S\,], \forall j \in \left[1, \lambda^i\right],\ i,j \in \mathbb{N},\ \ \exists\,(j_1,k_1),(j_2,k_2) \\ &\quad s.t.\ \ \left|c^i_{j_1,k_1} - \tilde{c}^i_{j_1,k_1}\right| \leq 1,\ \left|c^i_{j_2,k_2} - \tilde{c}^i_{j_2,k_2}\right| \leq 1 \ \text{ and } \ c^i_{j,k} = \tilde{c}^i_{j,k} \ \ \forall\,(j,k) \neq (j_1,k_1),\ (j_2,k_2)\,. \end{aligned} \tag{4.24}$$

this adjacency relation indicates that changing the route of a single car could affect the counts measurements reported by each sensor i in at most two different time steps. To make it more clear, suppose that Sara's car triggers a number of sensors everyday when she goes to her job in the morning. In order for differentially private protect her absence or her presence, we must consider changing her trajectory could change the reported counts of each sensor at 2 different times, one corresponding to one unit decrease in her usual trend of passing and the other corresponding to one unit increase in her new trend.
Now, denote by ϕ^i_k (4.10) and $\tilde{\phi^i_k}$, the two adjacent flow datasets. Then

$$\left\|\phi - \tilde{\phi}\right\|_2^2 = \sum_{k=0}^{\infty}\sum_{i=1}^{M}\left|\phi^i_k - \tilde{\phi}^i_k\right|^2 = \sum_{i=1}^{M}\sum_{k=0}^{\infty}\left|\phi^i_k - \tilde{\phi}^i_k\right|^2.$$

and for a sensor at the interface $i \to i+1$ the corresponding term is

$$\sum_{k=0}^{\infty} \left|\phi_k^i - \tilde{\phi}_k^i\right|^2 = \frac{1}{T^2(\lambda^i)^2} \sum_{k=0}^{\infty} \left| \sum_{j=0}^{\lambda_i} \left(c_{j,k}^i - \tilde{c}_{j,k}^i\right) \right|^2 .$$

Based on the adjacency relation (4.24), the counts $c_{j,k}^i$ and $\tilde{c}_{j,k}^i$ must be almost all identical, except for the fact that some vehicles A and B can cross the line of the sensor at different periods and in different lanes. Thus

$$\sum_{k=0}^{\infty} \left|\phi_k^i - \tilde{\phi}_k^i\right|^2 = \frac{1}{T^2\,(\lambda^i)^2} \sum_{k=0}^{\infty} \left| \sum_{j=0}^{\lambda_i} \left(c_{j,k}^i - \tilde{c}_{j,k}^i\right) \right|^2 \leq \frac{2}{T^2\,(\lambda^i)^2},$$

and finally,

$$\left\|\phi - \tilde{\phi}\right\|_2^2 = \sum_{i=1}^{M} \sum_{k=0}^{\infty} \left|\phi_k^i - \tilde{\phi}_k^i\right|^2 \leq \frac{2}{T^2} \sum_{i=1}^{M} \frac{1}{(\lambda^i)^2} : \Delta f^2. \tag{4.25}$$

Now based on Theorem 3 in chapter 2 , the mechanism publishing for each sensor the perturbed flow pseudo-measurements $\Phi_k^i = \phi_k^i + n_k^i$, where n_k^i are independent zero-mean white Gaussian noise signals with covariance $\kappa_{\delta,\epsilon}^2 \Delta f^2$, with Δf defined in (4.25) , is (ϵ, δ)-differentially private.

4.6.2 Density and Mode Measurements

The flow pseudo-measurements

$$\Phi_k^i = \phi_k^i + n_k^i \tag{4.26}$$

obtained from the Gaussian mechanism can be used in calculating the density pseudo-measurements, but this requires an additional mode estimate. However, estimating the mode of the traffic based on the count or the occupancy datasets and without sanitization, could compromise private information of individuals. In this section, we present our private mode measurement that is mainly based on the mode measurement model presented in Section (4.5). Similar to (4.24), the adjacency relation for the occupancies datasets of N user routes $O := \{o_{j,k}^i \ : \ k \geq 0, 1 \leq i \leq S, 0 \leq j \leq \lambda^i\}$ is

$$\begin{gathered} \text{for all } o,\ \tilde{o} \in O: \ \text{Adj}\,(o, \tilde{o}) \ \text{ iff } \ \forall k \geq 0, \ \geq 0, \ \ \forall i \in [1, S\,], \forall j \in \left[1, \lambda^i\right], \ i, j \in \mathbb{N}, \exists\,(j_1, k_1)\,, (j_2, k_2)\,, \\ \psi \in [0, 1] \ \ s.t. \ \left|o_{j_1,k_1}^i - \tilde{o}_{j_1,k_1}^i\right| \leq \psi, \ \left|o_{j_2,k_2}^i - \tilde{o}_{j_2,k_2}^i\right| \leq \psi \ \text{ and } \ o_{j,k}^i = \tilde{o}_{j,k}^i \ \ \forall (j,k) \neq (j_1,k_1)\,, \ (j_2,k_2)\,. \end{gathered} \tag{4.27}$$

For the occupancy data, we bound the allowed deviation on the reported occupancy, when we add or remove one vehicle. In effect, this means that we offer no privacy protection to vehicles that change the measured cumulative occupancy too much (or that change the average speed too much,

since the occupancy contribution of one car is proportional to the inverse of its velocity). In other words, the occupancy time due to a single vehicle, equal to

$$O_{car}T = \frac{l_{car}}{v_{car}} \tag{4.28}$$

with T the sampling period, and l_{car} the average car length. If changing the trajectory of a single vehicle changes the measured occupancy by ψ, the velocity of that vehicle must be

$$v_{car} = \frac{l_{car}}{\psi T}$$

and this velocity even for a 7 meters long car is $\frac{0.84}{\psi} km/h$, that potentially corresponds to a car sitting on a sensor line for big values of ψ .

On the other hand, considering the occupancy $0 \leq o^i_{j,k} \leq 1$, the adjacency relation(4.27) will result in a high sensitivity and the corresponding standard Gaussian perturbation mechanism leads to unreliable occupancy pseudo-measurements, especially when the number of single-loop detectors in the road increases. Instead of using the occupancy measurements directly to estimate the density, the strategy adopted here is to re-consider the mode pseudo-measurement model presented in Algorithm. 1 from a differential privacy perspective, i.e, checking out how model (4.22) behaves when the trajectory of a single vehicle changes. Thus, let's re-consider model(4.22)

$$M^i_k = \begin{cases} F & if \quad \mathbf{1}_{T_F - T_C}\left((\Phi^i_k, y^i_k)\right) = 1 \\ C & if \quad \mathbf{1}_{T_C - T_F}\left((\Phi^i_k, y^i_k)\right) = 1 \\ M^i_{k-r} & if \quad \left[\prod_{s=0}^{r-1} \mathbf{1}_{T_C \cap T_F}\left((\Phi^i_{k-s}, y^i_{k-s})\right)\right] \mathbf{1}_{\bar{T}_C \cup \bar{T}_F}\left((\Phi^i_{k-r}, y^i_{k-r})\right) = 1,\ r > 0 \end{cases} \tag{4.29}$$

here we replace flow measurements (4.10) by the flow pseudo-measurements Φ^i_k (4.26) . By changing the trajectory of a single vehicle, we have

$$\tilde{M}^i_k = \begin{cases} F & if \quad \mathbf{1}_{T_F - T_C}\left((\tilde{\Phi}^i_k, \tilde{y}^i_k)\right) = 1 \\ C & if \quad \mathbf{1}_{T_C - T_F}\left((\tilde{\Phi}^i_k, \tilde{y}^i_k)\right) = 1 \\ \tilde{M}^i_{k-r} & if \quad \left[\prod_{s=0}^{r-1} \mathbf{1}_{T_C \cap T_F}\left((\tilde{\Phi}^i_{k-s}, \tilde{y}^i_{k-s})\right)\right] \mathbf{1}_{\bar{T}_C \cup \bar{T}_F}\left((\tilde{\Phi}^i_{k-r}, \tilde{y}^i_{k-r}\right) = 1,\ r > 0 \end{cases} \tag{4.30}$$

defining $\tilde{y}^i_k - y^i_k = \Delta y^i_k$, $\tilde{\Phi}^i_k - \Phi^i_k = \Delta\Phi^i_k$, and according to the adjacency relations defined in (4.24),

(4.27), we have

$$\begin{aligned} &\forall i \in [1, S\,],\ \ \exists k_1,\ k_2\ \ s.t.\ \ \left|\Delta y^i_{k_1}\right| \le \frac{\psi}{g\lambda^i},\ \left|\Delta y^i_{k_2}\right| \le \frac{\psi}{g\lambda^i},\ \ \Delta y^i_k = 0\ \ \forall i \ne i_0, \\ &\left|\Delta \Phi^i_{k_1}\right| \le \frac{1}{T\lambda^i},\ \left|\Delta \Phi^i_{k_2}\right| \le \frac{1}{T\lambda^i},\quad \Delta \Phi^i_k = 0\ \ \forall i \ne i_0. \end{aligned} \tag{4.31}$$

Then, all the possible mode switching resulted from changing the trajectory of a single vehicle is shown in following lemma.

Lemma 12. *For the sets T_F and T_C defined in (4.17), (4.18), and all flows Φ^i_k, we have*

$$\begin{aligned} &\mathbf{1}_{T_F}\left((\Phi^i_k, y^i_k)\right)\mathbf{1}_{T_C}\left((\tilde{\Phi}^i_k, \tilde{y}^i_k)\right) = 0, \qquad \text{if}\ \ \Phi^i_k \notin [\alpha, q_{max}\,], \\ &\text{and}\ \ \mathbf{1}_{T_C}\left((\Phi^i_k, y^i_k)\right)\mathbf{1}_{T_F}\left((\tilde{\Phi}^i_k, \tilde{y}^i_k)\right) = 0, \end{aligned} \tag{4.32}$$

$$\text{where } \alpha = min\left\{\left[\frac{e^{-\zeta(g)}\left(\rho_{max} - \frac{1}{T\lambda^i w}\right) - \frac{\psi}{g\lambda^i}}{\frac{e^{\zeta(g)}}{v_f} + \frac{1}{e^{\zeta(g)}w}}, q_{max}\right], \left[\frac{e^{-\zeta(g)}\rho_{max} - \frac{e^{\zeta(g)}}{T\lambda^i v_f} - \frac{\psi}{g\lambda^i}}{\frac{e^{\zeta(g)}}{v_f} + \frac{1}{e^{\zeta(g)}w}}\right]\right\}.$$

Proof. Following Lemma 11 , and in view of (4.20) ,(4.21) ,

$$\begin{cases} \mathbf{1}_{T_F}\left((\Phi^i_k, y^i_k)\right)\mathbf{1}_{T_C}\left((\tilde{\phi}^i_k, \tilde{y}^i_k)\right) = 1 \quad \text{iff} \begin{cases} e^{-\zeta(g)}\dfrac{\Phi^i_k}{v_f} \le [e^{\zeta(g)}(\rho_{max} - \dfrac{\Phi^i_k + \Delta\Phi^i_k}{w})] - \Delta y^i_k \\ \text{and} \\ [e^{-\zeta(g)}(\rho_{max} - \dfrac{\Phi^i_k + \Delta\Phi^i_k}{w})] - \Delta y^i_k \le e^{\zeta(g)}\dfrac{\Phi^i_k}{v_f} \end{cases} \\ \mathbf{1}_{T_C}\left((\Phi^i_k, y^i_k)\right)\mathbf{1}_{T_F}\left((\tilde{\phi}^i_k, \tilde{y}^i_k)\right) = 1 \quad \text{iff} \begin{cases} e^{-\zeta(g)}(\rho_{max} - \dfrac{\Phi^i_k}{w}) \le [e^{\zeta(g)}\dfrac{\Phi^i_k + \Delta\Phi^i_k}{v_f}] - \Delta y^i_k \\ \text{and} \\ [e^{-\zeta(g)}\dfrac{\Phi^i_k + \Delta\Phi^i_k}{v_f}] - \Delta y^i_k \le e^{\zeta(g)}(\rho_{max} - \dfrac{\Phi^i_k}{w}) \end{cases} \end{cases} \tag{4.33}$$

that is to say

$$\begin{cases} F \to C \quad \text{if}\ \ \Phi^i_k \in A \\ C \to F \quad \text{if}\ \ \Phi^i_k \in B \end{cases}$$

where

$$A = \left[\frac{\left[e^{-\zeta(g)}\left(\rho_{max} - \frac{1}{T\lambda^i w}\right)\right] - \frac{\psi}{g\lambda^i}}{\frac{e^{\zeta(g)}}{v_f} + \frac{1}{e^{\zeta(g)}w}}, q_{max}\right],\quad B = \left[\frac{e^{-\zeta(g)}\rho_{max} - \frac{e^{\zeta(g)}}{T\lambda^i v_f} - \frac{\psi}{g\lambda^i}}{\frac{e^{\zeta(g)}}{v_f} + \frac{1}{e^{\zeta(g)}w}}, q_{max}\right]$$

where we fix the maximum of each interval at q_{max} in order to prevent any likelihood of privacy leakage. The minimums are also minimized over corresponding parameters based on (4.31). Finally, the proof is obtained by

$$\begin{aligned} &\mathbf{1}_{T_F}\left((\Phi_k^i, y_k^i)\right)\mathbf{1}_{T_C}\left((\tilde{\Phi}_k^i, \tilde{y}_k^i)\right) = 0, \qquad \text{if} \quad \Phi_k^i \notin A \cup B \\ &\text{and} \ \ \mathbf{1}_{T_C}\left((\Phi_k^i, y_k^i)\right)\mathbf{1}_{T_F}\left((\tilde{\Phi}_k^i, \tilde{y}_k^i)\right) = 0, \end{aligned} \tag{4.34}$$

□

The above lemma shows that for the flow pseudo-measurement $\Phi_k^i \leq \alpha$, changing the trajectory of a single vehicle does not affect the outcome of mode measurement model (4.29). Accordingly, we now develop our private mode measurement model. Based on Lemma. 12 , we first define the sets

$$PT_F = \left\{ \left(\Phi_k^i, y_k^i\right) \ : \ \left| \log\left[\frac{\phi_k^i}{v_f}\right] - \log\left[y_k^i\right] \right| \leq \zeta(g), \ \Phi_k^i \in [0, \alpha) \quad \forall i, k \right\}$$

$$PT_C = \left\{ \left(\phi_k^i, y_k^i\right) \ : \ \left| \log\left[\rho_{max} - \frac{\phi_k^i}{w}\right] - \log\left[y_k^i\right] \right| \leq \zeta(g), \ \Phi_k^i \in [0, \alpha) \ \ \forall i, k \right\}$$

corresponding to the pseudo-flow Φ_k^i satisfying our private truncation in free mode (F) or congested mode (C). Defining $\bar{PT}_C$ and $\bar{PT}_F$, the complement sets of PT_C and PT_F respectively, we could form the private mode measurement model as

$$M_k^i = \begin{cases} F & if \qquad \mathbf{1}_{PT_F - PT_C}\left((\Phi_k^i, y_k^i)\right) = 1 \\ C & if \qquad \mathbf{1}_{PT_C - PT_F}\left((\Phi_k^i, y_k^i)\right) = 1 \\ M_{k-r}^i & if \qquad \left[\prod_{s=0}^{r-1} \mathbf{1}_{PT_C \cap PT_F}\left((\Phi_{k-s}^i, y_{k-s}^i)\right)\right] \mathbf{1}_{\bar{PT}_C \cup \bar{PT}_F}\left((\Phi_{k-r}^i, y_{k-r}^i)\right) = 1, \ r > 0 \end{cases} \tag{4.35}$$

Similar to Fig. 4.5 , we could also divide the triangular fundamental diagram into two zones, we call here the **Private** zone and the **Non-Private** zone. For illustrative purposes, these zones are depicted with a green line and a red line on Fig. 4.5 . The Private zone for example, illustrates the flow interval for which the mode of the traffic can be estimated uniquely based on truncation (4.15), and also changing the trajectory of a single vehicle does not affect the mode estimation. The diagram is depicted for a four-lane road and the parameters related to the fundamental diagram and the model are $v_f = 65$ mph, $w = 11.6$ mph, $\rho_{max} = 193$ vehicles/mile/lane, $g = 20$ feet and $\zeta(g) = 0.51$. We also take $\psi = 0.25$ that is sufficiently large to protect the privacy of individuals, i.e, all the vehicles that cross the sensor line faster than $3km/h$, assuming vehicles with at least 7 meters long. Note that our model for the flows in Non-Private zone, estimates the mode of the

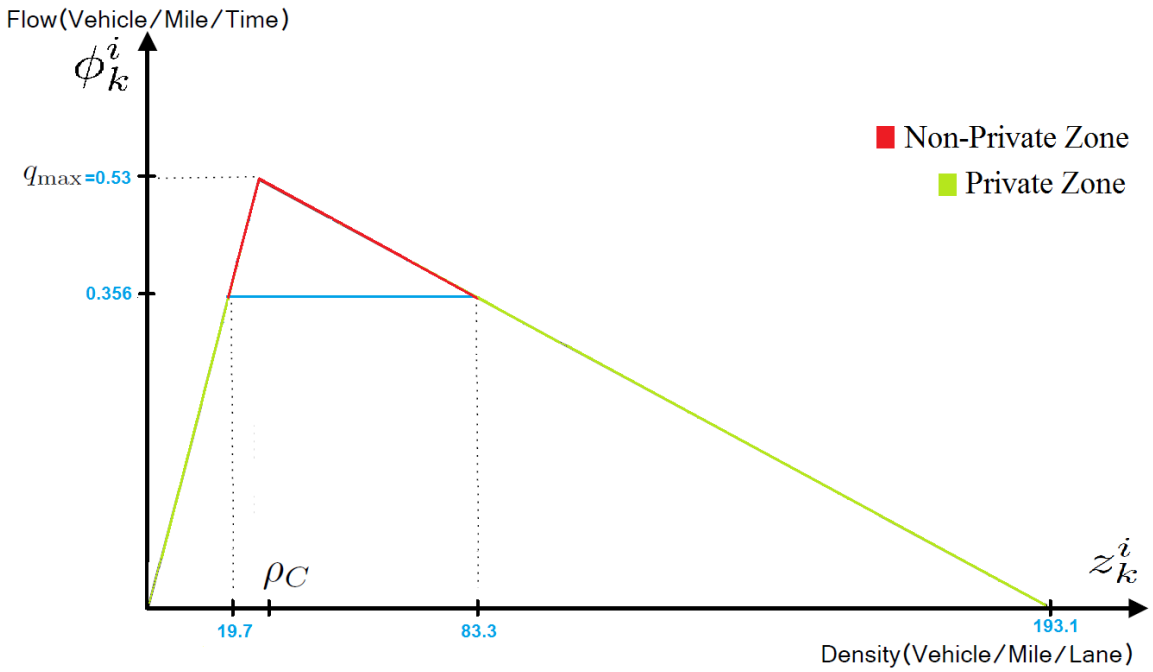


Figure 4.5 The mode of the traffic in private zone is resistant to the change of the trajectory of a single vehicle.

traffic at each sensor location based on the previous estimated mode for which the flow is in Private zone. Adopting this strategy minimizes the possibility of privacy leakage while it still provides a meaningful observation signal to specify the mode. Our model of obtaining the private density pseudo-measurements is demonstrated by Algorithm. 2 .

1- Perturb the flow measurements (4.10) to obtain the differentially private flow pseudo-measurements $\Phi_k^i = \phi_k^i + n_k^i$.
2- based on historical data choose base g-factor, e.g, 20 feet, and an upper-bound error $\zeta(g)$.
3- Set the maximum deviation ψ (4.28) in two adjacent occupancy data. Note that picking ψ too big spoils the mode estimation with the goal of protecting the privacy of too slow vehicles.
4- Specify the corresponding Private and Non-private zones based on Lemma (12).
5- $m_k^i = F/C$ based on the mode measurement model (4.35).
6- Filter m_k^i thorough the HMM filter to obtain the actual mode s_k^i used to invert the fundamental diagram.(see the discussion for HMM filter in Section 4.5)
7-

$$z_k^i = z_k^{i+1} = \begin{cases} \dfrac{\Phi_k^i}{v_f} & if \quad s_k^i = F \\ \rho_{max} - \dfrac{\Phi_k^i}{w} & if \quad s_k^i = C \end{cases} \tag{4.36}$$

Algorithm 2: Private density measurement

4.7 Traffic state Estimation

In this section, we present the overall architecture of our differentially private traffic state estimator, and illustrate its performance on the Mobile Century experiment dataset [36]. Fig. 4.6 illustrates the overall architecture of our privacy preserving traffic estimator. The extended Kalman filter (EKF) assimilates the dynamic traffic model (4.4) and the density pseudo-measurements z_k^i, obtained from the occupancy and count measurements.

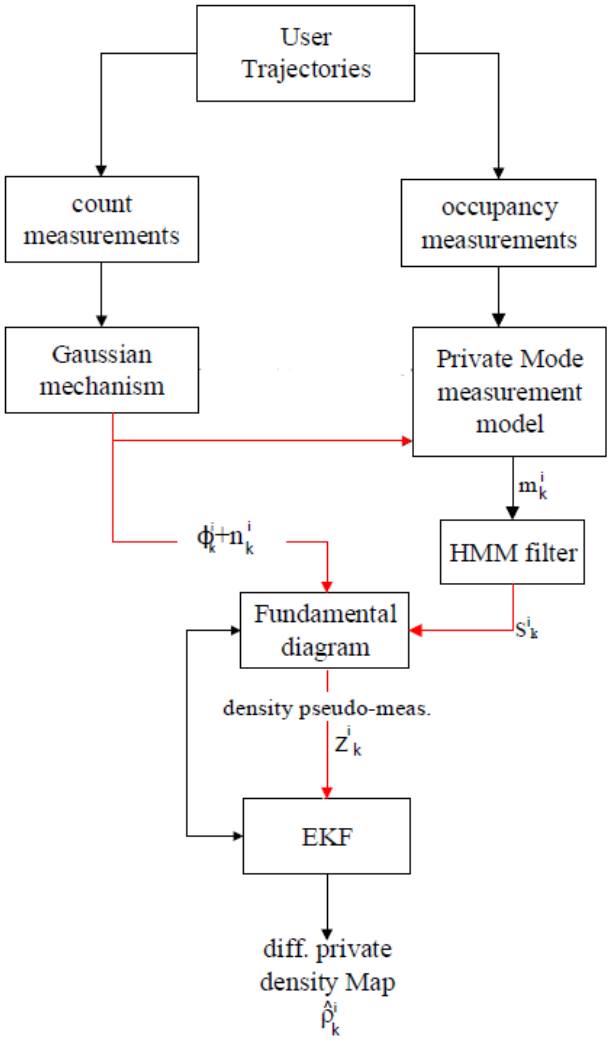


Figure 4.6 Architecture of our differentially private traffic estimator. The red arrows represent differentially private signals, i.e, perturbed flow pseudo-measurements from vehicle counts, and private mode estimate built from both counts and occupancy measurements

The differential privacy guarantee provided by this architecture is the sum of guarantees provided by the Gaussian mechanism and our private mode measurement model. Recalling Theorem 3 in chapter 2 , we now specify how much privacy guarantee is provided by our mode measurement model.

Lemma 13. *Consider* $D = \left\{\left(c^i_{1,k}, .., c^i_{\lambda^i,k}, o^i_{1,k}, .., o^i_{\lambda^i,k}\right) \ \ \forall i, k \right\}$*, consists of the count and occupancy datasets with adjacency relations defined in (4.24) , (4.27). Let d, d′ be two adjacent elements in D with K rows. Define* $E = \{E_i : E_i = \ e_1 \times e_2 \times ... \times e_K, \ e_i = [0, \alpha) \text{ or } [\alpha, q_{max}] \ \text{ for all } \ i = 1, ..., 2^K\}$*, with α defined in Lemma. 12 , specifying the private zone. Then for our private mode measurement mechanism M, and the flow pseudo measurement Φ, we have*

$$M(d) = M(d') \ \text{ if } \ \Phi(d), \Phi(d') \in E_i, \quad \forall d, \ d' \in D, \ \text{ and for all } i = 1, ...2^K. \tag{4.37}$$

Proof. For two pair of adjacent data d, d', if their pseudo-flows $\Phi(d)$, $\Phi(d')$ are in the same zone, the mode measurement model will result in identical outputs, because the model always estimate the mode with respect to the flows in private zone. Hence, the model automatically ignores any change in occupancy measurements resulted from adding or removing one vehicle. □

Theorem 14. *The private mode estimation mechanism defined in (4.35) is (ϵ, δ)-differentially private.*

Proof. Define $\chi = \{F, C\}^K$, for all $d, \ d' \in D$ and $s \in \chi$, we have

$$\mathbb{P}(M(d) \in s) = \sum_{i=1}^{2^k} \left[\mathbb{P}\left(M(d) \in s \mid \Phi(d) \in E_i\right) \mathbb{P}\left(\Phi(d) \in E_i\right)\right] = \sum_{i=1}^{2^k} \left[\mathbb{P}\left(M(d') \in s \mid \Phi(d') \in E_i\right) \mathbb{P}\left(\Phi(d) \in E_i\right)\right]$$

where the last equality resulted from (4.37) . The flow mechanism $\Phi(d) = \phi(d) + n$ is the output of a Gaussian mechanism and is (ϵ, δ)-differentially private, hence

$$\mathbb{P}\left(\Phi(d) \in E_i\right) = \frac{1}{(2\pi\sigma^2)^{k/2}} \int_{E_i} e^{-\frac{\|u - \phi(d')\|^2}{2\sigma^2}} e^{\frac{2(u - \phi(d'))^T(\phi(d) - \phi(d')) - \|\phi(d) - \phi(d')\|^2}{2\sigma^2}} du \leq e^{\epsilon}\mathbb{P}\left(\Phi(d') \in E_i\right) + \frac{1}{(2\pi\sigma^2)^{k/2}} \int_{E_i} \left[e^{-\frac{\|u - \phi(d)\|^2}{2\sigma^2}} \mathbf{1}\left\{2(u - \phi(d'))^T(\phi(d) - \phi(d')) \geq \|\phi(d) - \phi(d')\|^2 + 2\epsilon\sigma^2\right\}\right] du$$

and the last integral term defines a measure that is bounded by δ (we refer the reader for more details to proof of Theorem 3 in [16]). Define A be the flow area indicated by the indicator function, for the last integral we then have

$$\frac{1}{(2\pi\sigma^2)^{k/2}} \int_{E_i} \left[e^{\frac{\|u-\phi(d)\|^2}{2\sigma^2}} \mathbf{1}\left\{2(u-\phi(d'))^T(\phi(d)-\phi(d')) \geq \|\phi(d)-\phi(d')\|^2 + 2\epsilon\sigma^2\right\} \right] du$$
$$= \mathbb{P}(\Phi(d) \in [A \cap E_i]) = \mathbb{P}(\Phi(d) \in A)\, \mathbb{P}(\Phi(d) \in E_i \mid \Phi(d) \in A)$$

we know $\sigma^2 = \|\phi(d)-\phi(d')\|^2 \kappa^2_{\epsilon,\delta}$, by straightforward calculation we have

$$\mathbb{P}(\Phi(d) \in A) = \delta \;\text{ and }\; \mathbb{P}(\Phi(d) \in [A \cap E_i]) = \delta\, \mathbb{P}(\Phi(d) \in E_i \mid \Phi(d) \in A)$$

Hence

$$\mathbb{P}(M(d) \in s) \leq \sum_{i=1}^{2^k} \mathbb{P}\left(M(d') \in s \mid \Phi(d') \in E_i\right) \left[e^{\epsilon}\mathbb{P}\left(\Phi(d') \in E_i\right) + \delta\, \mathbb{P}(\Phi(d) \in E_i \mid \Phi(d) \in A)\right]$$
$$= e^{\epsilon}\mathbb{P}\left(M(d') \in s\right) + \delta \sum_{i=1}^{2^k} \mathbb{P}\left(M(d') \in s \mid \Phi(d') \in E_i\right) \mathbb{P}(\Phi(d) \in E_i \mid \Phi(d) \in A)$$
$$= e^{\epsilon}\mathbb{P}\left(M(d') \in s\right) + \delta \sum_{i=1}^{2^k} \mathbb{P}\left(M(d) \in s \mid \Phi(d) \in E_i\right) \mathbb{P}(\Phi(d) \in E_i \mid \Phi(d) \in A)$$

and the last sum is bounded by 1, because it is the summation for mutually exclusive events $\Phi(d) \in E_i$, conditioned on a single event $\Phi(d) \in A$. □

Finally, the differential privacy guarantee for the overall architecture based on Theorem 2 is the sum of $(2\epsilon, 2\delta)$ obtained from (ϵ, δ)-differential privacymode measurement and (ϵ, δ)-differential privacyflow pseudo-measurement.

4.7.1 Discussion

Fig. 4.7 and Fig. 4.8, show examples of $(\log(2), 0.05)$ and $(\log(4), 0.1)$-differential private maps respectively, based on our designed private traffic estimator. The complete map is built by using 10 out of the 27 sensors which placed at the 4-lane locations along I-880. Our result improves over the state of the art in terms of the privacy guarantee, and has immediate applications in providing privacy for traffic monitoring of long highways. The reliability of the map is also improved in the sense that the incorrect switching between the modes has mitigated significantly. Comparison of the three maps presented in Fig. 4.7 , 4.8 and 4.3 illustrates that we could provide strong

$(\log(4), 0.1)$, and even very strong $(\log(2), 0.05)$ privacy guarantees based on our approach and still do not have too much degradation in estimation performance. However, it is possible that our algorithm estimates the mode of the flows in non-private zone with error, mainly due to the delay introduced by our mode measurement model. For example, assume that the traffic flow increases and the road becomes congested. The mode of the traffic based on our private mode measurement model will be free (F) until the flow goes into the private zone again. In this case we could have up to 60 (vehicles/mile/lane) error in our density maps according to Fig. 4.5 . The upper-bound of these errors could be tightened by decreasing parameter ψ. For example, $\psi = 0.1$ can decrease the upper-bound of this errors to 40 units, but it will weaken the privacy guarantee at the same time. One effective idea to improve the mode measurement where the flow is in sensitive zone is

$$\text{for all } \Phi_k^i \in [\alpha, q_{max}] : \begin{cases} F & \text{if } \Phi_k^i - \Phi_{k-1}^i > 0 \\ C & \text{if } \Phi_k^i - \Phi_{k-1}^i < 0 \end{cases} \tag{4.38}$$

this model takes advantage of this fact that the dynamic of the flow is either decreasing or increasing corresponding to the congested (C) mode and free (F) mode respectively. However, unfortunately it seems unwieldy at this point to be used in a more advanced mechanism.

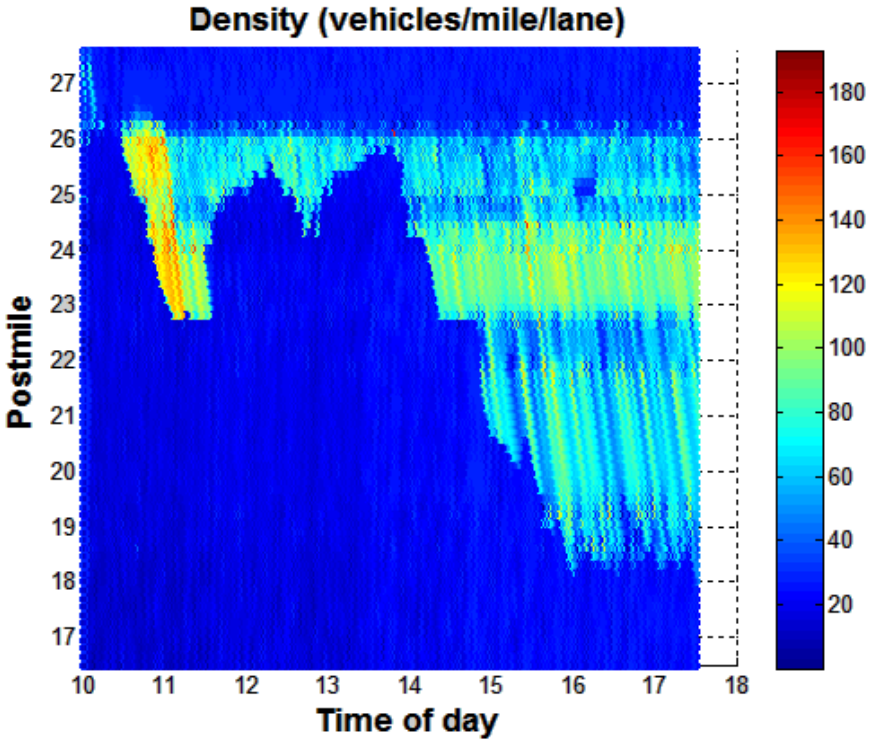


Figure 4.7 Real-time density map reconstruction with $(\log(2), 0.05)$- differential privacy guarantee presented based on our approach

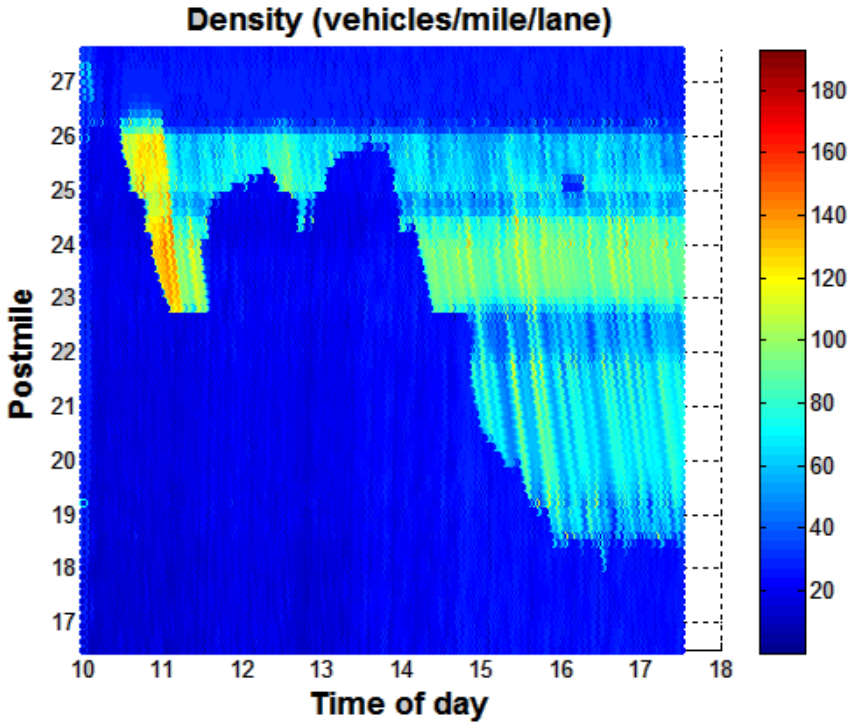


Figure 4.8 Real-time density map reconstruction with $(\log(4), 0.1)$- differential privacy guarantee presented based on our approach

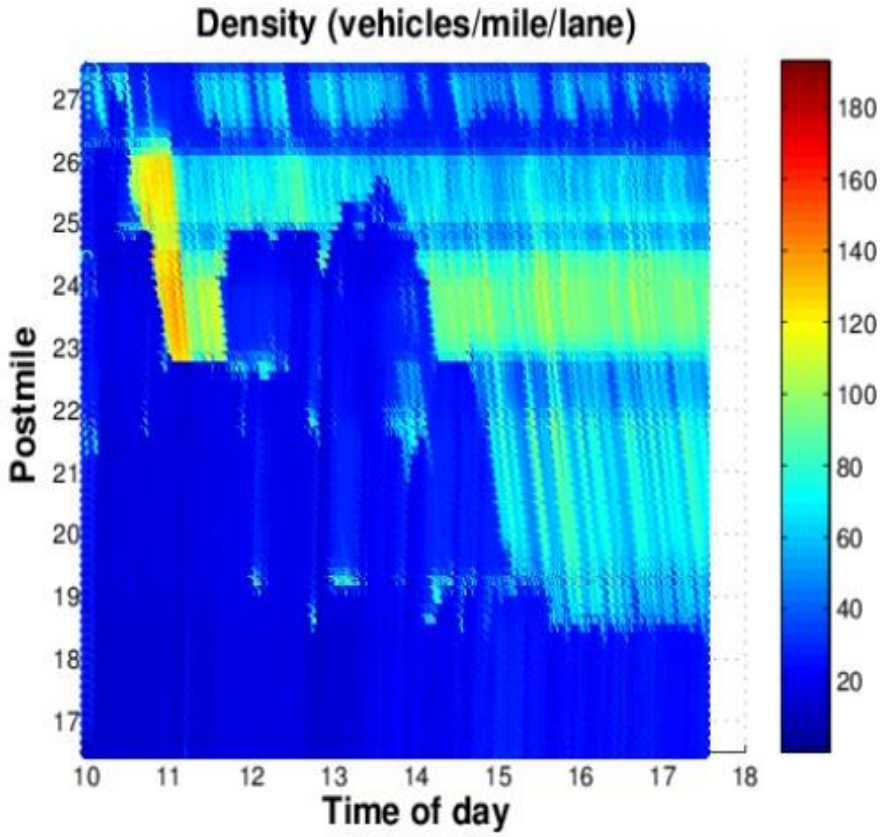


Figure 4.9 Real-time density map reconstruction with $(10 + \log(2), 0.05)$- differential privacy guarantee presented in [3]

CHAPITRE 5 CONCLUSION AND FUTURE WORK

In this Master's thesis we presented differentially private event stream filtering in a theoretical contexts for MIMO filters. The results were presented through rigorous mathematical proofs and based on them two important real world applications were studied. we have extended the ZFE mechanism of [16, 26] to the MIMO case. An optimal ZFE mechanism was obtained for the approximation of SIMO filters, and a suboptimal one considering only diagonal pre-filters was obtained for general MIMO filters. Next in Chapter 4, our differentially private traffic estimator was presented. Our results improves over state of the art and in particular the privacy preserving guarantee provided by our model improves very significantly over the previous design in [3]. Briefly, We presented

1. Differentially private event stream filtering in a theoretical contexts for MIMO filters.
2. Privacy preserving building monitoring
3. Privacy preserving traffic monitoring which describes techniques that can guarantee the differential privacy of individual users whose data is used to provide online estimation of the traffic state on a road section. In contrast to previously proposed privacy-preserving schemes for location-based services, we specifically target the release of aggregated quantities, such as effective traffic speed and density, and we rely on a macroscopic hydrodynamic model of the dynamics of these variables to provide sufficiently accurate estimators.

5.1 Future work

Future work includes developing MIMO mechanisms for situations where more information is available about the input signals, e.g., their second-order statistics, in which case one can improve on the ZFE mechanism [11]. Future work on designing a privacy preserving traffic estimator also includes improving the accuracy of mode measurements by applying an adaptive HMM filter and also studying the effects of increasing the number of sensors, and how this increasing affects the estimation.

RÉFÉRENCES

[1] A. Manta, "Literature survey on privacy preserving mechanisms for data publishing," technical report, Delft University of Technology, Delft, June/2013 2013.

[2] C. Wren, Y. Ivanov, D. Leigh, and J. Westhues, "The MERL motion detector dataset," Tech. Rep. TR2007-069, Mitsubishi Electric Research Laboratories, November 2007.

[3] J. Le Ny, A. Touati, and G. Pappas, "Real-time privacy-preserving model-based estimation of traffic flows," in *2014 ACM/IEEE International Conference on Cyber-Physical Systems (ICCPS)*, pp. 92–102, April 2014.

[4] M. Arrington, "AOL proudly releases massive amounts of private data," in *TechCrunch*, August 6 2006.

[5] M. Barbaro and T. Z. Jr., "A face is exposed forAOL searcher no. 4417749.," in *The New York Times*, August 9 2006.

[6] J. Bennett and S. Lanning, "The Netflix prize," in *In KDD Cup and Workshop in conjunction with KDD*, 2007.

[7] A. Korolova, *Protecting Privacy when Mining and Sharing User Data*. PhD thesis, Stanford University, 2012.

[8] A. Narayanan and V. Shmatikov, "Robust de-anonymization of large sparse datasets," in *2008. SP 2008. IEEE Symposium on Security and Privacy*, pp. 111–125, May 2008.

[9] L. Sweeney, "K-anonymity: A model for protecting privacy," *Int. J. Uncertain. Fuzziness Knowl.-Based Syst.*, vol. 10, pp. 557–570, Oct. 2002.

[10] C. Dwork, "Differential privacy," in *Proceedings of the 33rd International Colloquium on Automata, Languages and Programming (ICALP)*, vol. 4052 of *Lecture Notes in Computer Science*, Springer-Verlag, 2006.

[11] J. Le Ny, "On differentially private filtering for event streams," in *Proceedings of the Conference on Decision and Control*, (Florence, Italy), December 2013.

[12] V. Rastogi and S. Nath, "Differentially private aggregation of distributed time-series with transformation and encryption," in *Proceedings of the 2010 ACM SIGMOD International Conference on Management of Data*, SIGMOD '10, (New York, NY, USA), pp. 735–746, ACM, 2010.

[13] Y. D. Li, Z. Zhang, M. Winslett, and Y. Yang, "Compressive Mechanism: Utilizing Sparse Representation in Differential Privacy," *ArXiv e-prints*, July 2011.

[14] J. Le Ny and M. Mohammady, "Differentially private mimo filtering for event streams and spatio-temporal monitoring," in *Decision and Control (CDC), 2014 IEEE 53rd Annual Conference on*, pp. 2148–2153, Dec 2014.

[15] C. Dwork, "Differential privacy: A survey of results," in *Proceedings of the 5th International Conference on Theory and Applications of Models of Computation*, TAMC'08, (Berlin, Heidelberg), pp. 1–19, Springer-Verlag, 2008.

[16] J. Le Ny and G. Pappas, "Differentially private filtering," *IEEE Transactions on Automatic Control*, vol. 59, pp. 341–354, Feb 2014.

[17] S. Oh and P. Viswanath, "The composition theorem for differential privacy," *CoRR*, vol. abs/1311.0776, 2013.

[18] C. Dwork, F. McSherry, K. Nissim, and A. Smith, "Calibrating noise to sensitivity in private data analysis," in *Proceedings of the Third Conference on Theory of Cryptography*, TCC'06, (Berlin, Heidelberg), pp. 265–284, 2006.

[19] C. Dwork, K. Kenthapadi, F. McSherry, I. Mironov, and M. Naor, "Our data, ourselves: Privacy via distributed noise generation," in *Advances in Cryptology-EUROCRYPT 2006*, pp. 486–503, Springer, 2006.

[20] E. A. Lee, J. D. Kubiatowicz, J. M. Rabaey, A. L. Sangiovanni-Vincentelli, S. A. Seshia, J. Wawrzynek, D. Blaauw, P. Dutta, K. Fu, C. Guestrin, R. Jafari, D. Jones, V. Kumar, and R. Murray, "The Terraswarm Research Center," tech. rep., University of California at Berkeley, 2012.

[21] H. Chan and A. Perrig, "Security and privacy in sensor networks," *Computer*, vol. 36, pp. 103–105, Oct 2003.

[22] V. Rastogi and S. Nath, "Differentially private aggregation of distributed time-series with transformation and encryption," in *International Conference on Management of Data*, pp. 735–746, 2010.

[23] C. Dwork, M. Naor, T. Pitassi, and G. N. Rothblum, "Differential privacy under continual observations," in *Proceedings of the ACM Symposium on the Theory of Computing (STOC)*, (Cambridge, MA), June 2010.

[24] T.-H. H. Chan, E. Shi, and D. Song, "Private and continual release of statistics," *ACM Transactions on Information and System Security*, vol. 14, pp. 26:1–26:24, November 2011.

[25] J. Bolot, N. Fawaz, S. Muthukrishnan, A. Nikolov, and N. Taft, "Private decayed predicate sums on streams," in *Proceedings of the 16th International Conference on Database Theory*, ICDT '13, (New York, NY, USA), pp. 284–295, ACM, 2013.

[26] J. Le Ny and G. J. Pappas, "Differentially private filtering," in *Proceedings of the Conference on Decision and Control*, (Maui, HI), December 2012.

[27] D. H. Wilson and C. Atkeson, "Simultaneous tracking and activity recognition (STAR) using many anonymous, binary sensors," in *Pervasive Computing* (H.-W. Gellersen, R. Want, and A. Schmidt, eds.), vol. 3468 of *Lecture Notes in Computer Science*, pp. 62–79, Springer Berlin Heidelberg, 2005.

[28] J. Cao, Q. Xiao, G. Ghinita, N. Li, E. Bertino, and K.-L. Tan, "Efficient and accurate strategies for differentially-private sliding window queries," in *Proceedings of the International Conference on Extending Database Technology*, 2013.

[29] L. Ljung, *System Identification: Theory for the User*. Information and System Sciences, Prentice Hall, 1998.

[30] C. F. Daganzo, "The cell transmission model:a dynamic representation of highway traffic consistent with the hydrodynamic theory. transportation research part B: Methodological," 28(4):269-287,1994.

[31] M. Treiber and A. Kesting, "Traffic flow dynamics," pp. 66–80, Springer, 2013.

[32] M. J. Lighthill and G. B. Whitham, "On kinematic waves. II. A theory of traffic flow on long crowded roads," *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 229, no. 1178, pp. 317–345, 1955.

[33] P. I. Richards, "Shock waves on the highway," *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, February 1956.

[34] D. Simon, *Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches*. Wiley-Interscience, 2006.

[35] Z. Jia, C. Chen, B. Coifman, and P. Varaiya, "The PEMS algorithms for accurate, real-time estimates of g-factors and speeds from single-loop detectors," in *Intelligent Transportation Systems, 2001. Proceedings. 2001 IEEE*, pp. 536–541, 2001.

[36] J. C. Herrera, D. B. Work, R. Herring, X. J. Ban, Q. Jacobson, and A. M. Bayen, "Evaluation of traffic data obtained via GPS-enabled mobile phones: The mobile century field experiment," *Transportation Research Part C: Emerging Technologies*, vol. 18, no. 4, pp. 568 – 583, 2010.